Author:
Bonnington, Peter J

Title:
A systematic approach to failure mode characterisation in a complex sociotechnical context

A SYSTEMATIC APPROACH TO FAILURE MODE CHARACTERISATION IN A COMPLEX SOCIOTECHNICAL CONTEXT: A CASE STUDY IN DIESEL FUEL INJECTION SYSTEMS

Peter James Bonnington MEng (Hons) CEng MIMechE

A dissertation submitted to the University of Bristol in accordance with the requirements for award of the degree of Engineering Doctorate in Systems in the Faculty of Engineering

Faculty of Engineering, December 2018

70,638 Words

Abstract

Product Validation is an integral element of the New Product Development process: through identifying and correcting product failure modes, and demonstrating product reliability in line with customers' requirements, it drives many of the design and development activities of a business. If a products usage or failure mechanisms aren't fully characterised, the effectiveness of the Product Validation process is reduced. Failure Mode Characterisation is dependent on the knowledge held within a business and its capability to create and combine new knowledge in a timely, effective, and resource-sensitive manner, influencing the performance on the New Product Development process.

In the context of high-pressure diesel fuel injection systems, this thesis presents a case study of failure mode characterisation of seat wear of a hydraulic control valve, demonstrated to be a complex sociotechnical problem, which previously represented a source of partial uncertainty. This thesis asks whether a suitable method can be identified to characterise the usage variables that influence the seat wear, employing a multi-methodological approach. Expert Elicitation through the Delphi Method, and an expert panel with a combined 140 years of product knowledge, was used to capture and combine the existing knowledge of the failure mode. The resultant failure mode definition was then codified in a Causal Loop Diagram, with a total of 55 nodes, representing the complexity of the problem. Using Experimental Design methodology, a total of 60 samples were then tested for 1000 hours, identifying a usage variable as having a significant location effect on the failure mode, and demonstrating an injector performance metric that represents a suitable inference of degradation over time. Finally, regression modelling was used to generate a generalised model of the failure mode that can be used to predict the response over time for given usage conditions. This thesis concludes that the system operating pressure of the fuel system, that can vary with application, has a significant location effect on the wear of the hydraulic control valve, and can be inferred through injector performance degradation.

The main contribution to knowledge of this thesis is the characterisation of diesel fuel injector control valve seat wear with respect to the products usage, enabling accelerated tests to be designed, and robust designs solutions to be validated. In addition, a system model is presented that captures the variables associated with the design, manufacture, and usage of the fuel injector that influence the failure mode, ranked by perceived significance by the expert panel. Furthermore, the method utilised is presented as a generalised method for Failure Mode Characterisation in complex sociotechnical contexts. The method proposed represents a simple and useful model for practitioners, with emphasis on purposeful planning, knowledge creation and combination, reflection, and iteration.

Dedication and acknowledgements

Funding for this research was provided by the Engineering and Physical Sciences Research Council and Delphi Technologies. The author is employed by Delphi Technologies and supported by the Industrial Doctorate Centre in Systems Engineering & Management, Universities of Bristol and Bath, UK.

Professors Chris McMahon & Chris Brace: *For the last 14 years of undergraduate and postgraduate academic supervision, for providing consistently insightful and entertaining discussion, and for your invaluable contributions to this thesis.*

Professor Stuart Burgess: *For your inputs in the final stages of this thesis and for helping me to remember what the story was.*

Melanie Needham: *For hiring me all those years ago, giving me this opportunity, providing so much support, and being my second-best critic.*

My colleagues at Delphi Technologies: *For making the experimental work happen, providing additional sets of eyes for the results and findings, and, Robin Loftus, for turning each increasingly ridiculous request a reality.*

Friends and Family: *For bearing with me.*

Emma Bonnington: *For everything.*

Author's declaration

I declare that the work in this dissertation was carried out in accordance with the requirements of the University's Regulations and Code of Practice for Research Degree Programmes and that it has not been submitted for any other academic award. Except where indicated by specific reference in the text, the work is the candidate's own work. Work done in collaboration with, or with the assistance of, others, is indicated as such. Any views expressed in the dissertation are those of the author.

SIGNED: .... DATE:.........................

Table of Contents

List of Figures

List of Tables

Glossary and Acronyms

ADT	Accelerated Degradation Test
AF	Acceleration Factor
ALT	Accelerated Life Test
CV	Commercial Vehicles
DF	Deterioration Factor
DFR	Design for Reliability
DFSS	Design for Six Sigma
DOE	Design of Experiments
EPA	Environmental Protection Agency
EU	European Union
EUI	Electronic Unit Injector
EUP	Electronic Unit Pump
FAME	Fatty Acid Methyl Esters (Biodiesel)
FIE	Fuel Injection Equipment
FIS	Fuel Injection Systems
FMEA	Failure Mode and Effect Analysis
FMECA	Failure Mode, Effect, and Criticality Analysis
FMC	Failure Mode Characterisation
FRACAS	Failure Reporting, Analysis, and Corrective Action System
FTA	Fault Tree Analysis
HPRCS	High Pressure Common Rail System
ICE	Internal Combustion Engine
IDID	Internal Diesel Injector Deposits
ISC	In Service Conformity
KM	Knowledge Management
MDP	Minimum Drive Pulse, measured in microseconds (µs)
NCV	Nozzle Control Valve
NPD	New Product Development
NTE	Not to Exceed
NVH	Noise, Vibration, and Harshness
OBD	On Board Diagnostics

PCV	Pressure Control Valve
PDF	Probability Density Function
PEMS	Portable Emissions Measurement System
PSM	Problem Structuring Method
PV	Product Validation
QFD	Quality Function Deployment
RCA	Root Cause Analysis
SOP	Start of Production
TCO	Total Cost of Ownership
TI Drift	Trimmed Integral Drift, measured as area between two fuel gain curves (mm^3*µs)
UM Drift	Untrimmed Maximal Drift, measured in injected fuel quantity (mm^3)
WSD	Wear Scar Diameter, measured in microns (µm)

Euro VI	The current European Emissions Standard for Heavy Duty Vehicles, defining standards for CO_2 emissions, toxic emissions, and emission test cycles. OBD, noise, etc.
OEMs	Original Equipment Manufacturers: The businesses engaged in the design, development and manufacturers of heavy-duty vehicles
Tier 1	The businesses engaged in the design, development and manufacture of components and systems for direct supply to OEMs, representing the highest tier of the supply chain. Tier 1s will in-turn be supplied by lower tiers of the supply chain.

Chapter 1 Introduction

1.1 Background to Research

Businesses seeking competitive advantage strive to reduce the time to market for innovative and iterative product alike. However, those products must still meet the end-user's requirements with respect to both functionality and reliability. As such, to push a product to market without first demonstrating that it can meet those requirements can result in a product failing in the market place, ultimately to the detriment of the brand. Ensuring a product meets all customer requirements can require the application of significant product development resources, decreasing the potential profit from product sales, which in highly cost competitive markets such as the automotive industry, can place significant sociotechnical stresses on a New Product Development (NPD) process (Campean et al, 2013). When sufficient NPD resource is either not available, or reductions in potential profit are undesirable, businesses seek to improve their NPD process in order to achieve an improved return on their existing resources and capabilities.

Product Validation (PV) is a key element of the NPD process with regards to meeting the requirements of both the customer and the business and determines many of the resource requirements of the process. PV describes the process through which a product is demonstrated to meet all of the customers' requirements, including product reliability. Through identification and correction of product failure modes, and demonstrating product reliability, PV drives many of the design and development activities of a business, whilst identifying the required physical testing facilities. An effective PV process identifies required design changes early in the NPD process such as to reduce the costs associated with implementing and validating the changes. As such, it can be seen that PV is a critical process for businesses engaged in NPD, and interventions that can yield a reduction in the time and resource required to bring a product to market will represent a competitive advantage.

In order to facilitate informed design decisions in the NPD process, and to demonstrate a product's reliability, Accelerated Life Testing (ALT) is typically implemented in NPD. Through a detailed understanding of the product's design characteristics, its usage conditions, and possible failure modes, it is possible to design experiments that can accelerate the life cycle of a product in a controlled environment. This methodology can in itself reduce the resources required for an NPD programme, but requires significant knowledge to plan and implement successfully. If some of the required knowledge or empirical evidence is incomplete or unavailable, this introduces uncertainty, to the detriment of the effectiveness of ALT.

When managing uncertainty in NPD, 'Engineering Judgement', or more generally 'Expert Judgement', is a term commonly used to describe the application of subjective reasoning to reach a conclusion when the otherwise available data is unavailable or insufficient. In using such judgment, a business is drawing upon the expert knowledge held within its human resources, requiring knowledge to transfer across functional and geographic boundaries, with the potential for knowledge loss and misinterpretation.

The successful implementation of ALT is in part dependent on Failure Mode Characterisation (FMC). If uncertainty exists around the usage of a product, or in the mechanisms through which it might fail, then the effectiveness of the ALT methodology reduces to the detriment of the PV process. Effective PV from empirical testing, and initial field trials, is highly significant in NPD programmes with long life requirements (Doikin et al, 2018). FMC is dependent on the knowledge in the business, both codified and held within individual experts, and in the business' capability to create and combine new knowledge through its NPD process. It can therefore be seen that a business' ability to characterise failure modes in a timely, effective, and resource sensitive manner, has impact on the performance of both PV and the overall NPD process.

As automotive firms seek to introduce new products to the market with shorter NPD cycles, while reducing the cost of new product introduction, they do so on a roadmap towards Virtual Product Engineering (VPE), such as that published by the Automotive Council UK (Automotive Council UK, 2016). In such roadmaps, integrated Computer Aided Engineering (CAE) environments will ultimately be used to perform much of the PV in a simulation environment, but in order to do so, products, their applications, and their failure modes must be characterised. As such, any interventions that seek to improve the FMC capability of an organisation, while creating a new knowledge of the product, or a means for turning product and customer data into knowledge, contribute towards VPE in the automotive industry.

Bringing together a portfolio of elements from engineering design and soft systems methodology, and applying systems thinking to overcome the sociotechnical complexity associated with NPD programmes, this thesis presents a structured method for FMC through an industrial case study. In the context of high-pressure diesel Fuel Injection Systems (FIS), this thesis presents a case study in the characterisation of a failure mode that had not previously been fully characterised in the form of seat wear of a fast-acting hydraulic control valve of a Diesel Fuel Injector (DFI).

The additional outcome of this thesis is a generalised FMC method presents a means for reducing uncertainty in NPD programmes, through a structured framework of expert elicitation, knowledge transfer and combination, and experimental design. The method presented is suggested to enable better-informed decisions to be made earlier in the NPD process, while providing product and failure mode knowledge that can be codified into product design and analysis tools.

1.2 Introducing the sociotechnical complexity of this research

This case study presented in this thesis is embedded within Delphi Technologies, a global Automotive Tier 1 supplier of powertrain technologies. More specifically, the author is a member of the Commercial Vehicle (CV) FIS function of Delphi Technologies, an organisation with three technical centres and two manufacturing facilities in the UK, with additional engineering and manufacturing resource from the global Delphi Technologies increasingly utilised. The CV FIS function designs, develops, and manufactures FIS for the global market, influenced by the wider business, its shareholders, legislators, and trends in both the industry and its customers. Multidisciplinary teams, each with their own worldviews and experiences and often spread across different physical locations, operate in matrix organisational structures on product platforms with industry leading technology. Suppliers and sub-contractors are used to provide both finished components and processes alike. While design specifications and tolerances are tightly specified and measured, variations in material removal processes, heat treatments, coatings, and assembly processes compound such that part-to-part variation in the FIS products supplied to customers will always exist.

As such, the CV FIS business represents a complex sociotechnical context, with uncertainty and variation propagation through the product design, develop, and manufacturing process. The following series of figures will show how the complexity of the system increases while alluding to the geographical dispersion of the system, and how variation is manufacture and usage continue to increase as the products flow to usage in application at a high level of abstraction. Figure 1 provides an overview of the CV FIS business as a sub-system, highlighting the different actors within the system, the internal and externals influencers, and elements of variation.

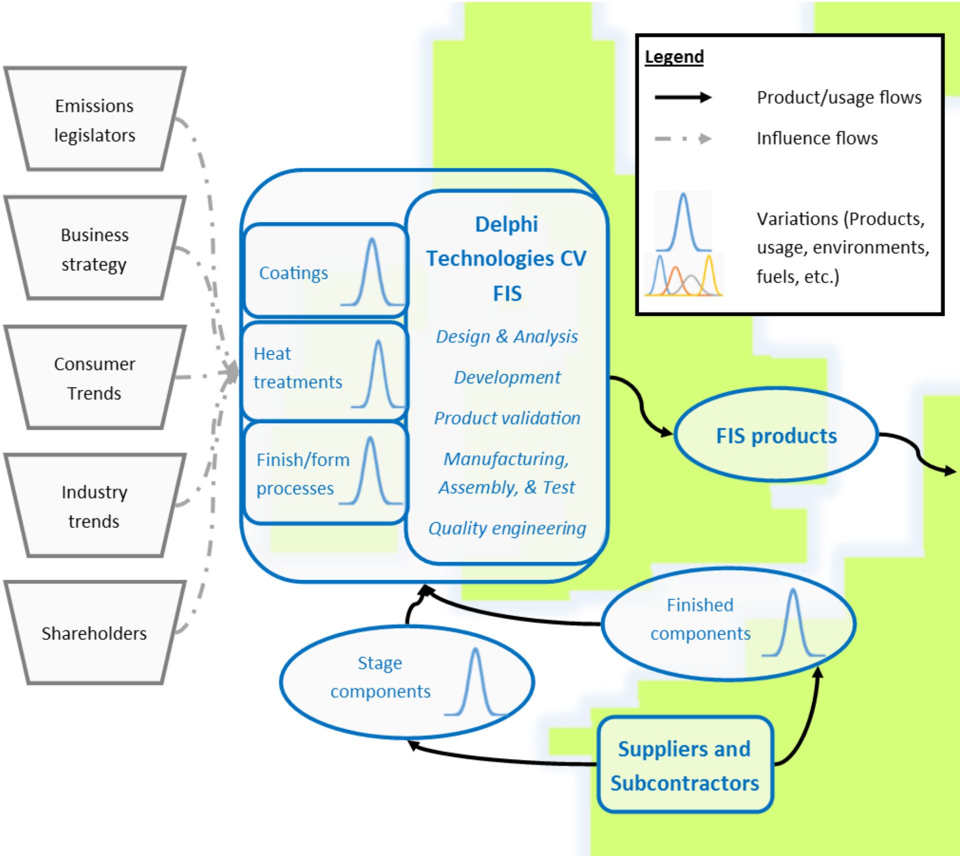


Figure 1: CV FIS business as a system

Figure 2 then shows the next sub-system in the form of Delphi Technologies' direct customers. In the CV market, the customers are generally Original Equipment Manufacturers (OEMs) who manufacture whole vehicles and/or standalone engines for sale to other OEMs and coach builders. The OEMs are influenced by a similar set of influencers as Delphi Technologies and will also be manifested in multi-disciplinary engineering teams. Each OEM will cater its products for a number of anticipated application duty cycles, each one representing an individual use distribution. Different performance and after-treatment strategies will also be considered, each influencing requirements of the FIS. Finally, each OEM will typically manufacture a number of different vehicle variations, some elements of which will in turn influence the use conditions of the FIS. As such, the engines and vehicles an OEM provides to its own customers will represent a number of discrete use distributions with respect to the FIS.

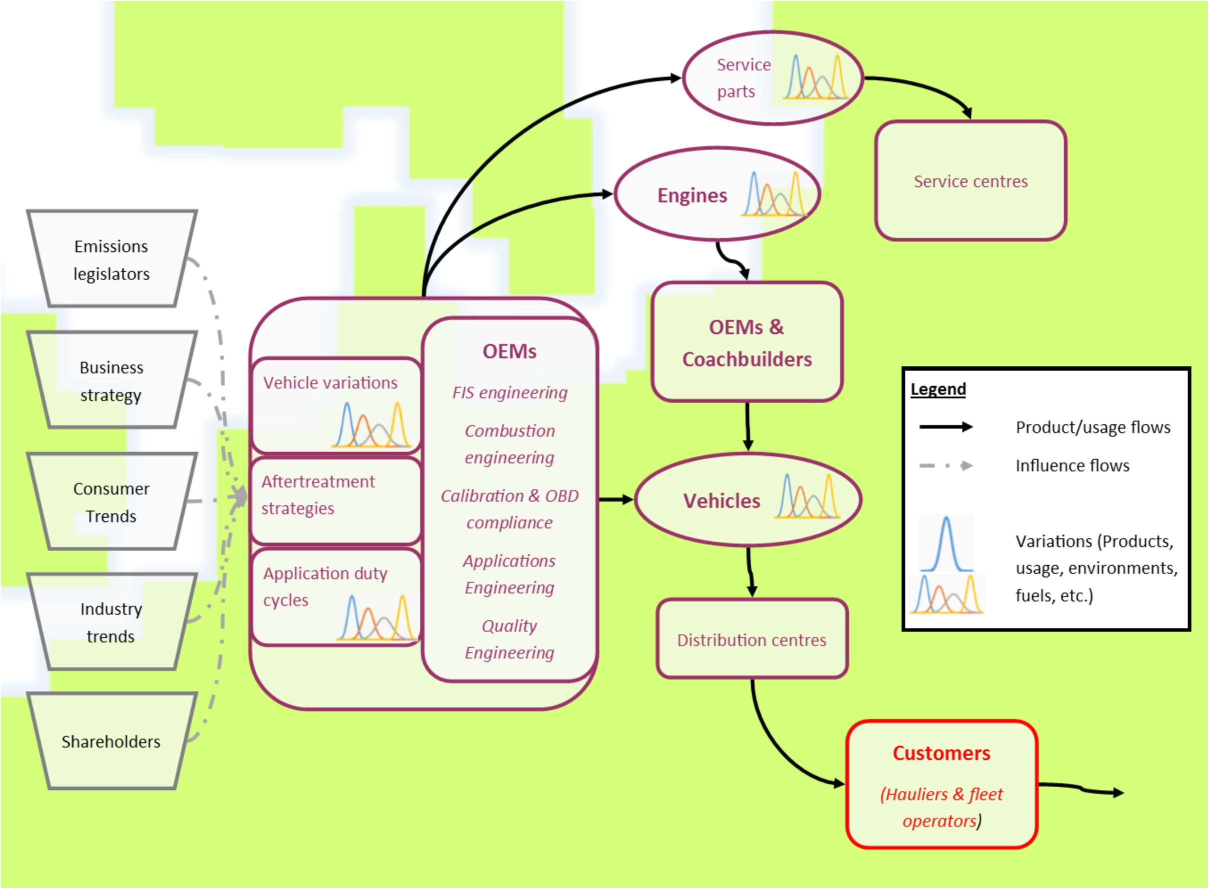


Figure 2: Original Equipment Manufactures as a system

Figure 3 then visualises the next sub-system, representing the end-users of the vehicles. The OEMs and coachbuilders market vehicles to the end customers who are typically hauliers and fleet operators. Through those customers, the combinations of end-user vehicle and operator(s) are formed, with driver behaviours representing additional variations. The fleets will either rely on the OEM's service centres or their own service centres to maintain the vehicles and FIS, using a combination of OEM service items and aftermarket products, with varying degrees of service schedule adherence. The particularities associated with a specific customer's actual duty cycles will combine with driver behaviours, road conditions, and environment conditions to influences the end-use vehicle duty cycles, resulting in a significant number of discrete use distribution for factors affecting FIS life.

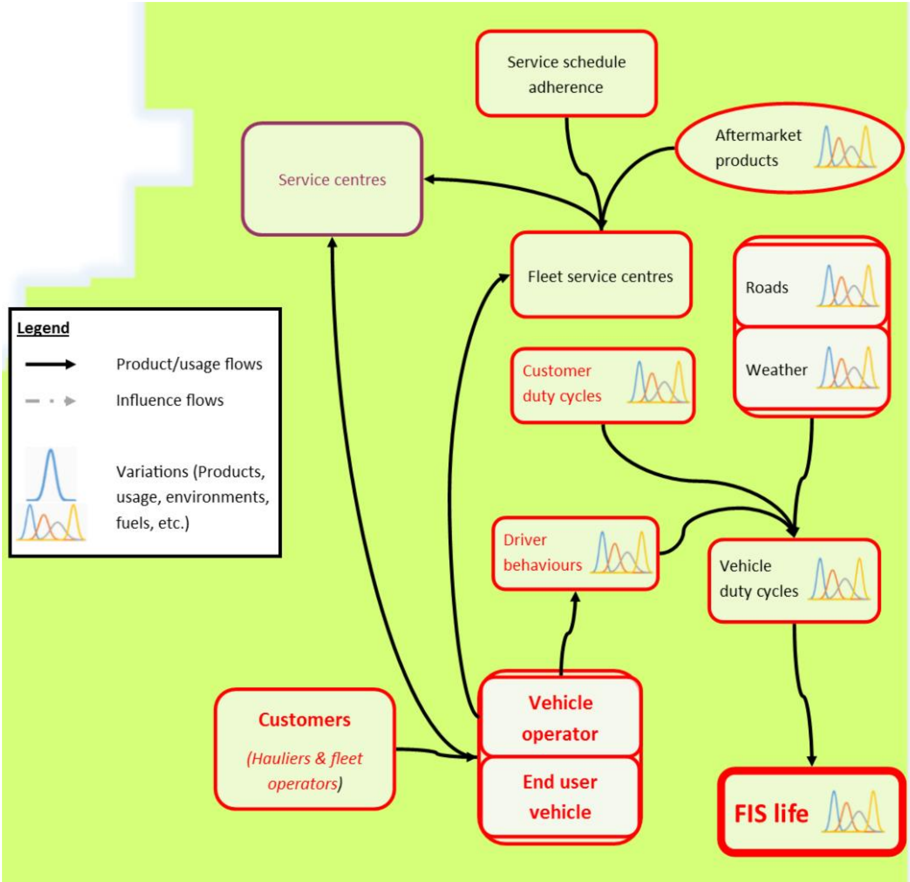


Figure 3: End Users as a system

The final consideration in the overall system for FIS is the fuel used in application, as visualised in Figure 4. The fuel suppliers are influenced legislation and market trends, developing seasonal fuel blends, that vary additionally with biodiesel blend ratio and feedstock, resulting in a number of discrete distributions of whole sale diesel fuels. Those fuels are then subject to the influence of additional discrete variations in the form of additives and non-desirable contaminants. Fuel distributors then supply that wholesale fuel to filling stations, with additional opportunity for the introduction of additives and contaminants. Finally, the end-user operator fills the vehicle with fuel at filling centres, with additional opportunities for the introduction of additives and contaminants. As a result, the fuels used in the end-user vehicle can vary significantly from tank to tank and can be exhibit a wide range in properties relevant to FIS life.

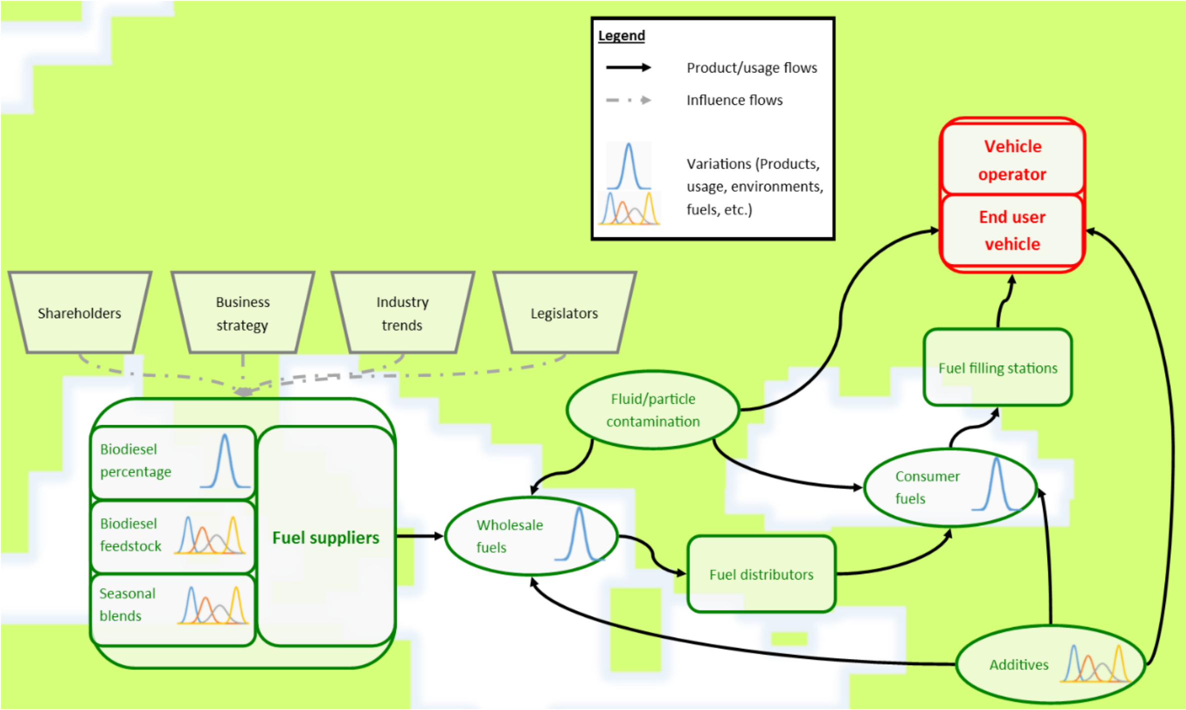


Figure 4: Fuel supply as a system

This series of systems combine to a singular, geographical distributed complex system that describes the factors that influence the life of an FIS product, as visualised in Figure 4. In this system, the variation associated with manufacturing of the products then propagates through to OEMs, who then produce a multitude of vehicle and powertrain variations, each placing different stresses on the FIS system in operation. The end users then introduce additional variation in the form driver behaviour, service adherence, and drive environments.

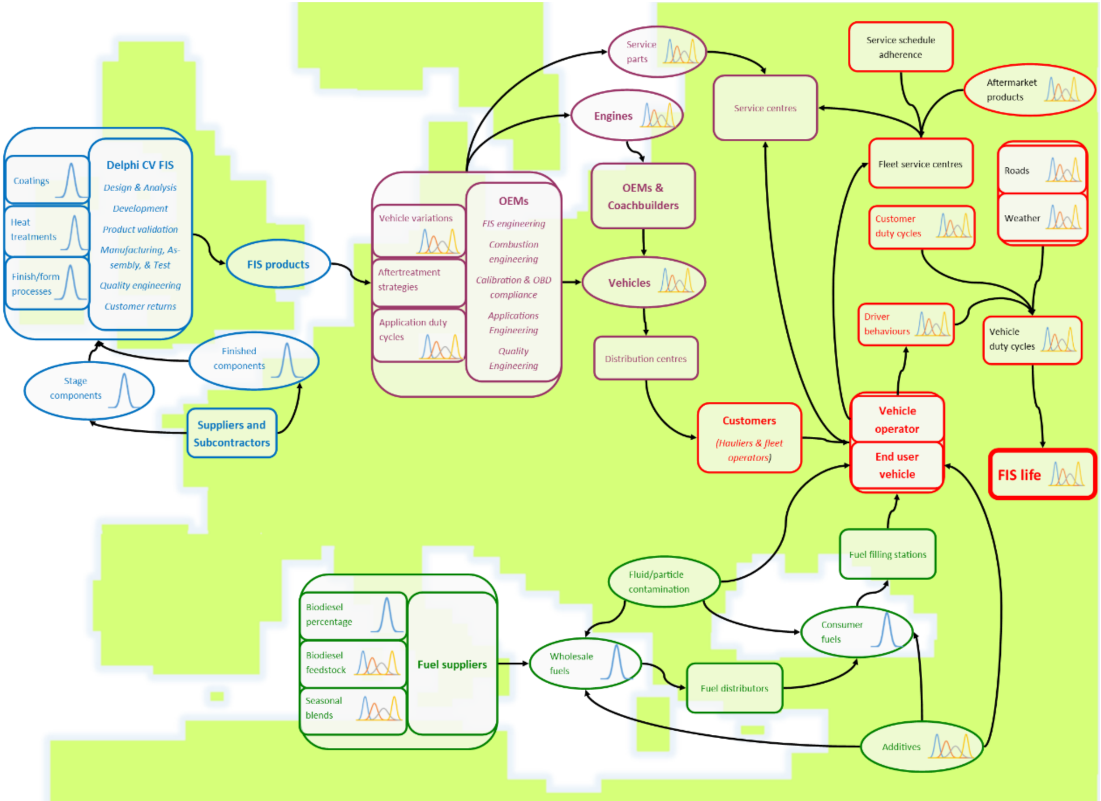


Figure 5: Visualisation of system complexity

These are combined with variations in fuel specification, cleanliness, and additives in the field, which have further influence over the life of FIS products. This compound complexity can be represented in Equation 1:

$$Product\ variation\ in\ manufacture\ +\ vehicle\ variations\ +\ duty\ cycles\ +\ fuel\ variations$$
$$=\ FIS\ life \tag{1}$$

Such a complex sociotechnical system results in the possibility that Delphi Technologies, or more generally any Tier 1 supplier, could be removed from many of the factors that influence product life by several degrees of separation. An example of this separation is visualised in Figure 6. In this example, the vehicle operator experiences a problem, either through the on-board diagnostic system or otherwise, that would be reported to the fleet service centre, or directly to the OEM service centre. The service centre would then diagnose the problem and remove any faulty FIS products, which would then be returned for processing through the OEM's warranty and quality departments. After possible involvement of the OEM engineering functions, the FIS would then be returned to Delphi Technologies for the fault to be confirmed and diagnosed. The product engineering functions of Delphi Technologies would then be consulted as appropriate in order to assist in any containment and corrective actions required. As such, it can be seen that the product engineering functions of Delphi Technologies, or any Tier 1 supplier, could be removed by approximately 5 degrees of separation, and a significant time delay, from the problems that may occur in application. When combined with the many variations that combine to influence FIS life, including uncertainty of the specific fuel(s) used, this fact can result in significant difficulties in accurately understanding the factors that influence product life for existing and indeed future products, and can contribute in a product failing to meet customer requirements.

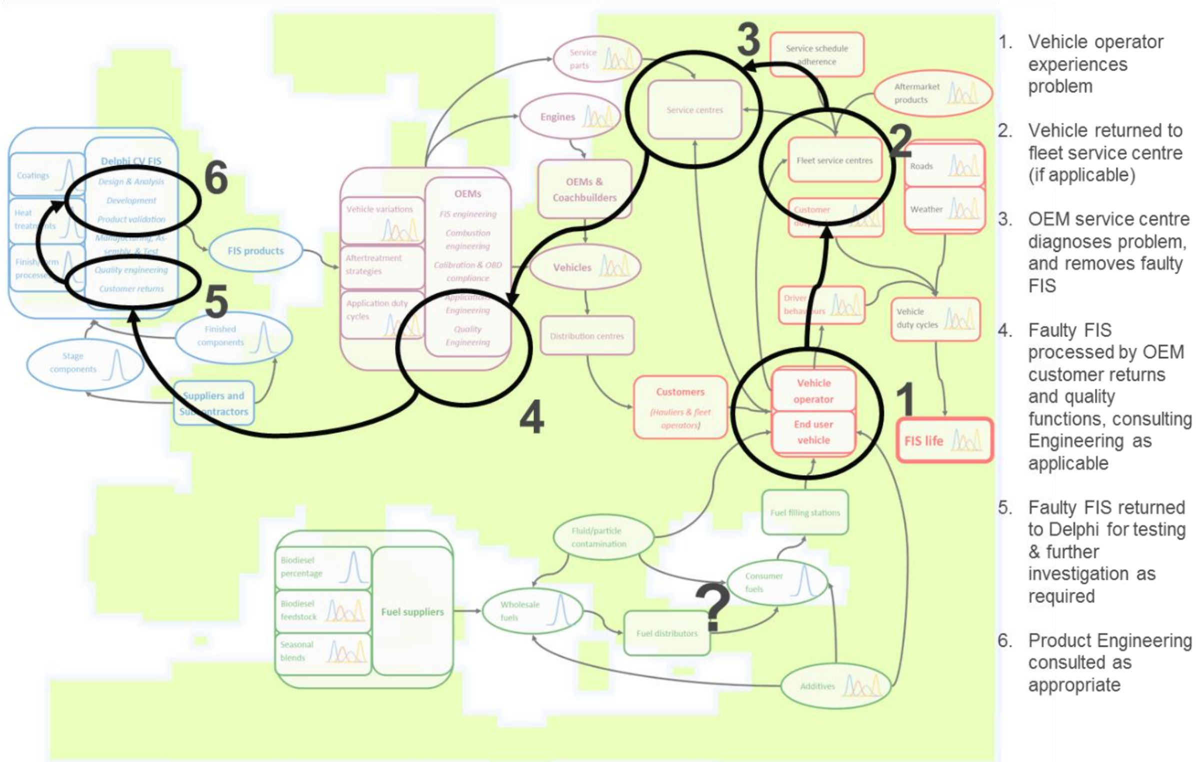


Figure 6: Visualisation of separation and lag between end-user and product engineering

Existing product design methods, including Quality Function Deployment and Failure Mode and Effects Analysis, exist such that with a robust understanding of the end users' requirements and usage profiles, organisations would be able to design robust products. However, when sociotechnical systems are suitably complex, as presented in this case study, the requisite knowledge of requirements and usage conditions are subject to significant uncertainty. When product failures are then experienced, that uncertainty influences the ability of an organisation to successfully characterise the failure mode, to the detriment of the PV process.

1.3 Research Hypothesis and Questions

With a view to forming a viable research hypothesis, this thesis takes the opportunity to review the individual elements of systems engineering and experimental design methodologies typically employed in industry. Through the application of systems thinking to the problem presented, this thesis is structured around designed interventions in an industrial case study, with the overall objective being to demonstrate the research against the following research hypothesis:

In a complex sociotechnical organisation engaged in the development of Fuel Injection Systems, can a systematic method for failure mode characterisation be developed that creates new knowledge of a failure mode subject to previous, unsuccessful investigations, and can that method be generalised for organisations engaged in New Product Development.

The generalised FMC method can be described using Soft Systems Methodology through the following root definition:

A method to improve failure mode characterisation, through resource sensitive expert elicitation and designed experiments, in order to reduce the resources required to validate new products.

To develop specific research objectives, and to provide a structure to this thesis, the following research questions were identified during the course of the development of the FMC method;

RQ1: How can expert judgement relevant to DFI control valve seat wear be elicited in an effective, and resource sensitive manner?

RQ2: Can a suitable modelling technique be identified to codify the results of the expert elicitation as a useful boundary object in the NPD process?

RQ3: Does structured empirical investigation represent an effective and resource sensitive means for identifying the significant effects and interactions associated with the usage variables identified through expert elicitation?

RQ4: Can a generalised model be developed from the empirical results that adequately describes the failure mode with respect to existing knowledge, while serving as the basis for further investigations?

RQ5: Can this method for FMC be generalised as a resource sensitive method for complex sociotechnical contexts?

1.4 Statement of contributions

Owing to the author's role as an employed Research Engineer embedded in industry, this thesis contributes both generalisable contributions to knowledge, and immediate industrial impact.

1.4.1 Contribution to Knowledge

The contribution to knowledge of this thesis is twofold. Firstly, a case study of FMC of an electronically controlled, hydraulic control valve with respect to its usage is presented in the context of high-pressure diesel FIS. In addition, a system model is presented that captures the variables associated with the design, manufacture, and usage of the fuel injector that influence the failure mode, ranked by perceived significance by the expert panel. In doing so, this research also represents a case study in the application of the Delphi Method, Causal Loop Diagrams, and Fraction Factorial experimental designs.

The second contribution to knowledge of this thesis is the generalised systematic method for failure mode characterisation in complex sociotechnical contexts derived from the case study. The method is presented in a spiral model, and represents a simple and useful model for practitioners, with emphasis on planning, the combination of new and existing knowledge, reflection, and iteration. This method is suggested to enable automotive firms to move towards VPE.

1.4.2 Industrial Impact

This research has been conducted with the support of Delphi Technologies, for whom the author is employed as a practitioner within the Diesel FIS business line. This research impacts industry in two ways, the first of which is a direct result of the case study presented in this thesis concerning a previously uncharacterised failure mode. Through the characterisation of the usage variables that influence DFI control valve seat wear, and the development of a generalised regression model that describes failure mode progression over time for different conditions, Delphi Technologies has been able to develop controlled accelerated tests for validation of robust design solutions, representing an immediate and lasting impact.

Secondly, the FMC method developed in this thesis has been codified into the Delphi Technology product validation process. This capability improvement is intended enable knowledge about key failure modes to be made available earlier in the NPD process in a resource sensitive manner, reducing the resources required to validate new products while maintaining or improving field reliability and customer satisfaction. A longitudinal study, outside the scope of this thesis, will be conducted to demonstrate this improvement, while the method itself has been demonstrated as transferable outside of the Diesel FIS business line.

1.5 Thesis structure

This thesis presents a case study of failure mode characterisation in diesel FIS, suggesting the method used could be generalizable for firms engaged in NPD in a complex sociotechnical context. The structure is based on the following:

Chapter 2: Research Methodology discusses how this thesis is developed around a systematic approach that draws from multiple disciplines, resulting in the use of multi-methodology in a plurality of world views. The research methodologies chosen for each research question are introduced, alongside a brief review of the alternative methodologies, before then discussing the implications of the research design with regards to both ethics and validity.

Chapter 3: Diesel Fuel Injection Systems provides an overview of the background, practitioner literature, & state of the art associated with the design, operation, and failure of the DFI that form the basis of the case study presented in this thesis. The operational characteristics of a typical Delphi Technologies Euro VI DFI are discussed, while an overview of the market drivers and typical product characteristics and properties employed to meet them are presented. The fundamental properties of diesel fuels that impact the operation of DFI will then be introduced. Finally, the implications of the technological trajectory of DFI on will be presented, with an overview of industry-wide failure modes of increasing severity and prevalence.

Chapter 4: Review of the Literature presents the background associated with the three main academic domains associated with this research: expert judgement; experimental design; and reliability engineering. An overview of the background of knowledge in the NPD process will be presented, with specific discussion around expert judgement and expert elicitation, focusing on the challenges presented in the NPD process. The background of the Experimental Design methodology will then be presented, providing an overview of the methods available, methodological limitations, and applications in industry. Finally, the background of the Reliability Engineering techniques employed to validate the functionality and reliability of products will be presented, focusing on the role of engineering judgement in the NPD process, and the role of accelerated testing in generating product knowledge.

Chapter 5: Existing Technologies, Processes, Tools, & Methods Within Delphi Technologies provides a brief sociotechnical contextualisation of this research. The methods and metrics typically associated with injector performance characterisation is presented, alongside the methods employed by Delphi Technologies in the validation of new products, including FMC and ALT. The role of expert judgement in the NPD process, along with methods for expert elicitation, will be presented with regards to the academic literature presented in Chapter 4, focusing on the role of knowledge in the process and the propagation of uncertainties.

Chapter 6: NCV Seat Wear - Eliciting and Codifying Expert Judgement introduces the case study that forms the basis for this research. An overview of the case study will be presented, both in terms of the failure mode itself, and the industrial motivations in this research. An application of Expert Elicitation via the Delphi Method is then presented as the approach for generating a group-led definition of the failure mode, and the variables that influence it. The Delphi Study also codifies the expert judgement of the significance of each variable on the severity and progression of the failure mode. System Modelling is then presented as the approach for codifying the elicited judgement into a purposeful system model in the form of a Causal Loop Diagrams, while visualising the complexity of the failure mode through a total of 55 nodes and 78 interactions.

Chapter 7: NCV Seat wear – Characterisation through empirical testing presents the second stage of the case study, in which the codified expert judgement is used as the inputs into a structured empirical investigation using Experimental Design Methodology. The variables identified in Chapter 6 are reviewed for their suitability for use as design factors, and appropriate response factors are identified and developed. A pair of iterative experiments are presented, serving to inform the final experimental design and variable selection. A Fractional Factorial design is presented as the methodology for this case study, with three design factors, each tested at two levels. An overview of the resulting data is then presented, alongside the analysis of the Experimental Design main effects and interactions. A centre point is then introduced into the experimental design in order to assess the linearity of the most significant effect.

Chapter 8: NCV Seat Wear – Modelling the failure mode presents the final stage of the case study, in which the results of the empirical investigation is used as the inputs in a regression modelling process. The performance change over time data associated with each test is used to fit a series of non-linear regression models, with additional performance characterisation intervals at low hours used to better characterise the failure mode progression. A generalised model of the failure mode is then developed in terms of both time, and levels of the significant usage variable. The generalised model is then validated using alternative empirical results, before then being used to compare different application duty cycles, and tests cycles.

Chapter 9: Discussion of Results and Generalisation of the Method presents a reflection of both the case study itself, and the methods used. The results expert elicitation study and resulting system model will be reflected upon, and the suitability of the approach as part of the FMC method will be discussed. The results of the experimental study will then be discussed with respect to both new knowledge and remaining knowledge gaps. A reflection on the knowledge gained through the regression modelling process, and its role within the FMC method will also be presented. Finally, the approach employed in this thesis will be presented as a generalisable method for FMC, alongside a discussion of how different FMC methods can be measured relative to each other.

Chapter 10: Conclusions presents a summary of this thesis, reflecting on the outcomes of the research against the research hypothesis, and specific research questions. Furthermore, the individual research contributions are identified.

If the reader is familiar with FIS technology, and the typical tools and methods employed by practitioners in NPD, then it is suggested that Chapters 3 and 5 could be excluded from the reading order respectively.

1.6 Summary

This Chapter has introduced the background and motivations for this research and has presented the research hypothesis along with a series to research questions to be answered in this thesis. The thesis itself has been presented as a case study of FMC in industry, with a generalisable method suggested as being effective in resource sensitive contexts. The next Chapter will describe the methodology and framework of the case study.

Chapter 2 Research Methodology

This Chapter will provide an overview of the research methodology associated with this research, demonstrating that messy, sociotechnical problems can require a multitude of methodologies. The research methods associated with each research question will be presented, followed by a discussion of research Ethics and Quality in the context of this thesis.

2.1 The Need for Multi-Methodology

Mingers (2011) defines wicked, messy problems as those which are not well defined, involve a multitude of interested parties and their associated multitude of worldviews, are characterised by uncertainties, and involve crossing the boundaries between social, and technical engineering systems. Such problems represent complexity, differentiated from complication. In this research, the problem context is that of product validation of FIS, which while technically complicated, and representative of a hard engineering problem, of itself does not represent complexity. As the research is embedded in a soft system, Delphi Technologies, it represents a complex, sociotechnical problem. Sociotechnical systems describe the overlapping social and technical systems typically associated with complex engineering projects, where the two systems co-exist, and their interaction is part of the problem context.

In the context of this research, and in the more general case of multi-disciplinary organisations engaged in the NPD of engineering systems, multiple and diverse stakeholders are involved, from engineers to accountants, manufacturing operators to the end users, each from different disciplines, and each with different perspectives, and each with different worldviews.

Problem Structuring represents a structured framework for dealing with complexity, and Problem Structuring Methods (PSMs) describe non-mathematical, but rigorous methods for doing so. PSMs allow for the multitude of worldviews associated with complex problems, and encourage engagement with stakeholders, while being tolerant to uncertainty (Rosenhead, 1996) (Mingers and Rosenhead, 1996)

Kuhn (1962) describes research paradigms as the set of beliefs and arguments shared between researchers that relate to the understanding and solving of problems. A number of different research paradigms have been identified, and can be characterised through their ontology, epistemology, and methodology, relating to the underlying research assumptions (Huba, 1990).

The ontological assumption concerns the researcher's view on the nature of reality. One may assume that reality is objective, and absolute, regardless of the viewer, or one may assume that reality is subjective, with a multitude of realities associated with different individuals based on their own worldviews.

The epistemological assumption concerns the researcher's view on knowledge, and the acknowledgement of knowledge as valid. One may consider reality as that which can be measured, with a focus on validated tools to achieve that measurement, or one could consider instead consider attitudes and opinions, interpreted by the researcher through interaction.

The methodological assumption concerns the researcher's views on the process of research. One could consider the process as being deductive, where theory is tested by the data, or one could consider the process as being inductive, where data is used to form a theory.

A number of research paradigms have been classified, including Positivism, Phemonology, and Pragmatism, and exist on a continuum (Mingers & Brocklesby, 1997). It is possible for a research hypothesis to be addressed using any one of a number of different paradigms, and in the event of complex problems, utilisation of a mixture of different paradigms can be necessary for dealing with individual research questions. Another factor influencing the selection of research paradigms is the personal disposition of the researcher and key stakeholders with regards to their ontological and epistemological worldviews.

2.2 Research Design

As discussed in Chapter 1, this research draws from several different academic domains. Several Research Questions have been identified, each one associated differently with those domains, and while perhaps a single paradigm could be used to approach each, that might not yield the most appropriate results, nor do so in such a way that matched the pre-disposition of the audience.

For this research, representing a complex sociotechnical problem, approaching research questions with tools from different domains, the most appropriate PSM is that of Multimethodology, where the research method applied to each Research Question will exist on the continuum between paradigm extremes.

The research paradigm that represents the closest direct analogy is that of *Pragmatism*. The ontological assumption associated with Pragmatism is that reality is in a state of constant flux based on our interventions, to be debated and interpreted by those with different worldviews. The Pragmatic paradigm centres around the research question and applies multiple methodologies to understanding that problem (Creswell, 2003). As such, research methods and data collection tools, both quantitative and qualitative, are employed when they are most likely to provide insight into the research question, regardless of philosophical paradigms they may be typically associated with (McKenzie & Knipe, 2006).

2.3 Research Design for each Research Question

RQ1: How can expert judgement relevant to DFI control valve seat wear be elicited in an effective, and resource sensitive manner?

RQ2: Can a suitable modelling technique be identified to codify the results of the expert elicitation as a useful boundary object in the NPD process?

While it would be possible to answer these research questions with methods from either end of the continuum of paradigms, a positivistic survey or a phenomenological ethnographic study as examples, a pragmatic approach is considered. In the context of this research, while the codified results of numerous, fragmented prior investigations were available and could suitable for a positivistic, deductive approach, there was a desire among the stakeholders to 'start over', engaging instead with the individual experts within the business to combine their uncodified knowledge. The knowledge would then be captured in a system model that could itself go on to inform research from either end of the continuum of paradigms. As such, with regards to this research question, the process is inductive, with the resulting characterisation being grounded in rich, qualitative data derived from the expert elicitation process, supported where available, by pre-existing codified knowledge. Chapter 6 will discuss the selection of an appropriate research strategy in the form of the Delphi Method for Expert Elicitation. Furthermore, the model building process will again be inductive, with a model being developed to make sense of the expert judgement, structuring it in such a way that further theories could be formulated to explain the system's interactions and dynamic behaviours. That model could also then be compared to existing positivistic models and be used to inform further investigations. Chapter 6 will also discuss the selection of a modelling method in the form of Causal Loop Diagrams. The inductive approach selected for both of these research questions will further the understanding of the complex sociotechnical context of this case study.

RQ3: Does structured empirical investigation represent an effective and resource sensitive means for identifying the significant effects and interactions associated with the usage variables identified through expert elicitation?

RQ4: Can a generalised model be developed from the empirical results that adequately describes the failure mode with respect to existing knowledge, while serving as the basis for further investigations?

For these research questions, a more positivistic paradigm is employed, combining experimentation and modelling in deductive theory testing. Chapter 7 outlines how the results of the Expert Elicitation and Systems Modelling elements identified a number of variables suitable for further investigation, and the process in which the variables suitable for structured empirical investigation within the scope of the of this case study were identified. Selection of an appropriate experimental methodology is then discussed, resulting in the use of Experimental Design to characterise the main effects and interactions associated with a number of variables. In Chapter 8, a modelling strategy is then employed to further characterise the failure mode through Regression Modelling. As such, this process represents elements of both deductive and abductive reasoning, where the observations generated through the process will be used to generate the most likely hypothesis that describes the nature of the failure mode. This positivistic approach was selected in part in order to produce results, in the form of statistically significant numerical models, that accommodated the ontological and epistemological worldviews of the research stakeholders.

RQ5: Can this method for FMC be generalised as a resource sensitive method for complex sociotechnical contexts?

The approach to this research question represented a more phenomenological one, with the overall process itself considered as the data, and the generalised method itself representing the theory. The method employed in this case study is discussed in Chapter 9, with lessons learnt presented from reflections associated with each element. The overall method itself is then discussed and presented using an adaptation of a spiral model, grounded in the case study. Methods for longitudinal assessment of the effectiveness of the model are also discussed.

Table 1 summarises the research design for each of the 5 research questions identified.

Research Question	Research Paradigm	Research Approach	Research Strategy	Time Horizon	Chapter
1	Phenomenology	Inductive	Grounded Theory – Delphi Method	Cross Sectional	6
2	Phenomenology	Inductive	Grounded Theory – Causal Loop Diagrams	Cross Sectional	6
3	Positivism	Abductive	Experiment – Experimental Design	Cross Sectional	7
4	Positivism	Abductive	Modelling – Regression modelling	Cross Sectional	8
5	Phenomenology	Inductive	Grounded Theory	Longitudinal	9

Table 1: Research design for each research question

2.4 Research Ethics

With this research embedded in an industrial context, research ethics are a significant consideration. Collis and Hussey (Colin & Hussey, 2003) identify ethical concerns for research in both natural and social sciences, including confidentiality; informed consent; dignity and publications. For this research, confidentiality is of consideration with regards to the intellectual property of Delphi Technologies, and as such, any measurements, tolerances, and quantitative data presented in this thesis will be anonymised. As elements of this research involve significant interaction with stakeholders and other actors within the organization, informed consent is also considered, with the appropriate approvals gained from the organisation's senior management and consent obtained from any individuals engaged through the research methods identified. Elements of the research allow individuals to present evidence to support any disagreements in expert judgement they may have with their peers; retaining the dignity of the individuals will be considered such as to avoid any potential embarrassment. Finally, all publications related to this research, both internal to the organisation, and those used externally, are required to be reviewed and approved by management and legal staff as appropriate, to protect both IP and the individuals involved in the research.

2.5 Research quality

Research quality can be described as the product of the reliability, validity, and generalisability of the research. The reliability of research concerns how well the research findings could be replicated. With regards to research questions 1 & 2 of this thesis, only a cross section of the experts within Delphi Technologies were used, and the results of the elicitation process were documented. As such, it would be possible to both conduct a repeat of the method using a different sample of experts, and for another researcher to audit the data in which the resulting knowledge is grounded. With respects to research questions 3 & 4, no non-standard parts, fuels, or processes were used, and tests were conducted and analysed using controlled & automated routines. As such, it would similarly be possible to both repeat the process using a new, equivalent sample population, and to audit the results and calculations.

Research validity describes to what extent do the findings of research accurately represent what is really happening. In order to improve the level of certainty, all qualitative results associated with this thesis are based on the consensus of a panel of experts, or results from stakeholder engagement. Furthermore, the quantitative results of this thesis can all have statistical confidence levels applied.

Finally, research generalisability describes to what extent can the results be applied to cases, settings, or situations beyond those associated with a particular case study. While this research is embedded in a specific case study, generalisations will be drawn to firms engaged in NPD in a complex sociotechnical context, and the method used will be presented as a generalised approach for FMC in such contexts.

2.6 Summary

This Chapter has presented the overall research paradigm, and specific research strategies to be used in this research. This research will use a Pragmatic research paradigm, using the research methods most appropriate to understand the central research problem in the social context of multiple world views, resulting in a multimethodological approach. The ethical considerations associated with this research have also been presented. The next Chapter will provide the background theory associated with the specific technological context of this research, through the market drivers that drive the development of FIS, the resultant product properties and characteristics embodied in the design of current state of the art, and their implications on the product’s potential robustness in application.

Chapter 3 Diesel Fuel Injection Systems: Operational overview and failure modes

3.1 Introduction

This Chapter provides an overview of the background theory & practice associated with the design, operation of the diesel fuel injection systems that provides the context to the transformation process described in this thesis.

Section 3.2 will provide an overview of the operation of Diesel Fuel Injectors, introducing the technology that forms the basis for this thesis. The following sections will then introduce the market drivers for FIS technology (§3.3), the associated product properties and characteristics employed to meet those drivers (§3.4), and the properties of diesel fuels that are of significance to the operation and robustness of Diesel FIS (§3.5).

Section 3.6 will then discuss the potential implications of the socio-technological trajectory of Diesel FIS in relation to its influence on the robustness of the products. An overview of the most significant failure modes that are generally exhibited in application will then be presented (3.7).

The Chapter will close with conclusions concerning the relationship between product technology, fuel properties, and the potential robustness of the product in application.

Figure 7 provides a system diagram that contextualises this Chapter, demonstrating the flow of FIS products into engines/vehicles, and ultimately to the end users. The flows of influence are also visualised, demonstrating the multitude of market drivers that influence FIS technology.

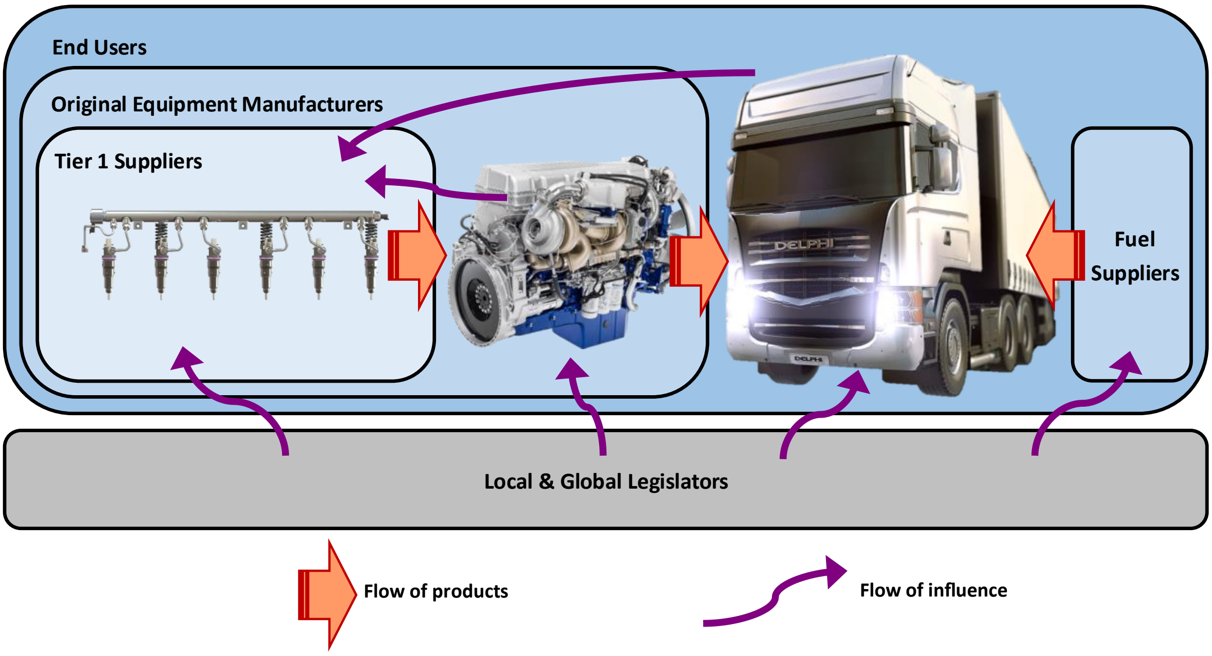


Figure 7: System diagram of flows of product and influence

3.2 Fundamental operation overview of diesel fuel injectors

This section will provide an overview of the fundamental operation of the "F2" family of Euro VI Diesel Fuel Injectors (DFI) developed by Delphi Technologies to provide focal context to this research. While the focus will be on the fuel injectors and the specific technology strategies employed to govern their operation, references will be made to the remainder of the common rail fuel system. Further detail will be provided about the electronically controlled hydro-mechanical valve employed to provide accurate and repeatable delivery of fuel.

3.2.1 Diesel Fuel Injector technology

Figure 8 shows a typical Delphi Technologies remote pump HPCRS fuel system, where 6 DFIs are connected to a common rail, itself supplied by a remote high-pressure pump. The system is controlled by an Electronic Control Unit (ECU) through the use of a Rail Pressure Sensor (RPS), and an electronically controlled metering valve on the pump. Furthermore, an electronic Pressure Control Valve (PCV) is integrated into the common rail to increase the capabilities of the rail pressure control system.

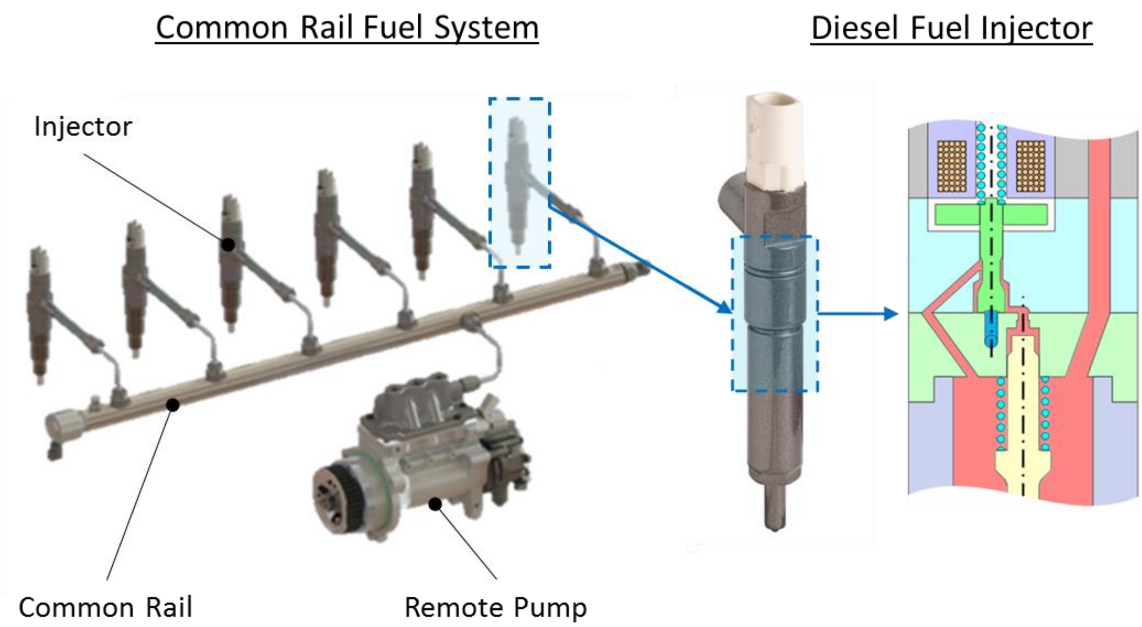


Figure 8: Typical remote pump HPCRS system

Figure 9 then shows the fundamental layout of a common rail DFI, identifying the key electromagnetic, mechanical, and hydraulic components.

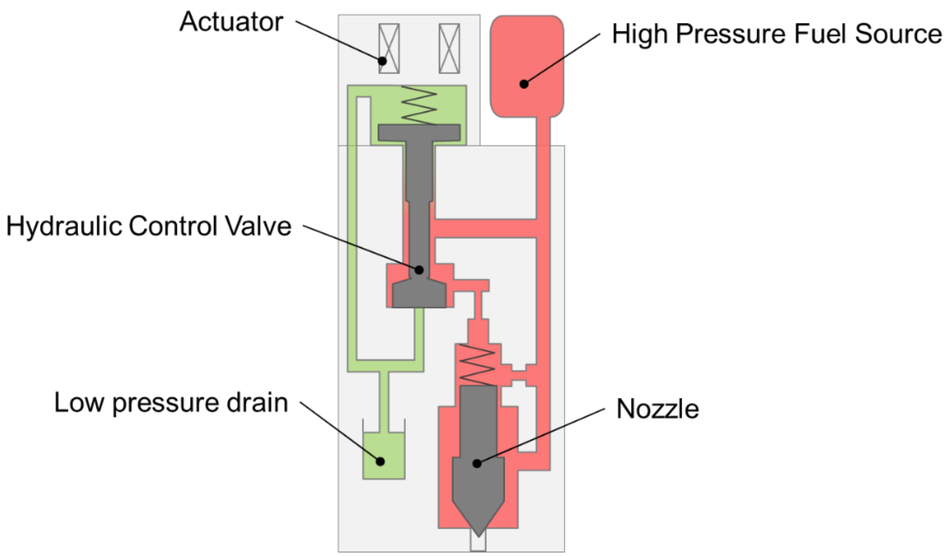


Figure 9: Fundamental layout of a common rail DFI

3.2.2 Delphi Technologies Euro VI Diesel Fuel Injectors

Figure 10 shows a simplified cross section for a typical DFI produced by Delphi Technologies, highlighting the two key assemblies in the form of the Actuator Assembly and the Nozzle Assembly. The Actuator Assembly consists of an electromagnetically actuated hydraulic control valve that controls the nozzle needle motion. The Nozzle Assembly then consists of a nozzle needle operating in a large volume of high-pressure fuel, controlling injection flow through the nozzle spray holes.

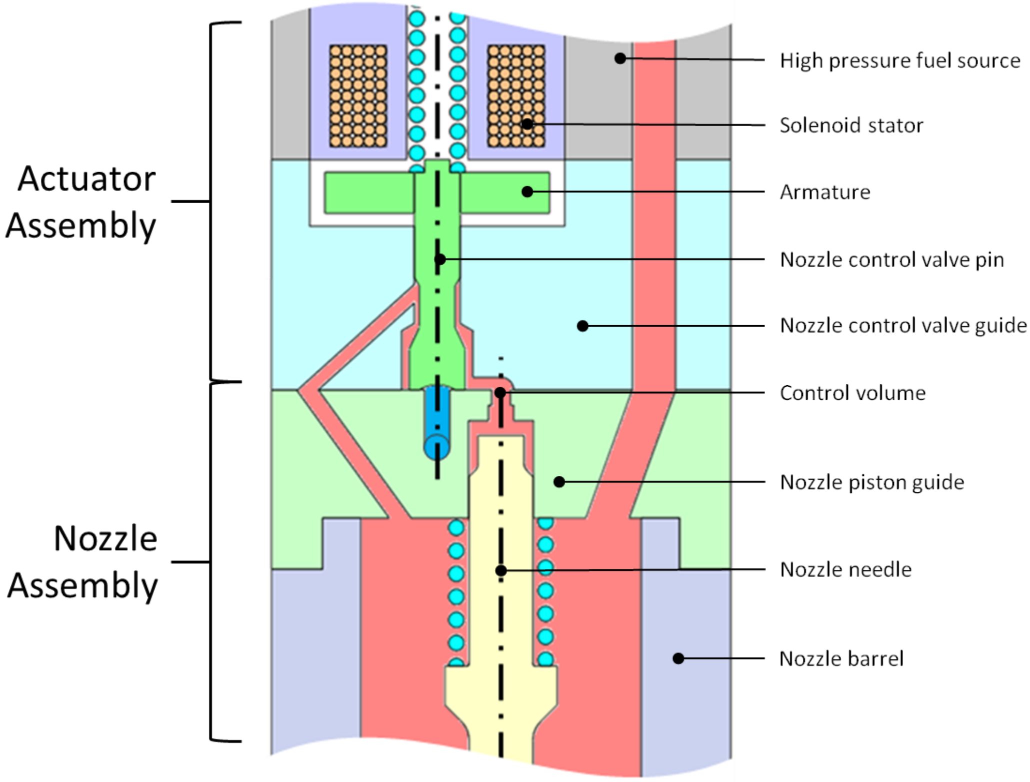


Figure 10: Simplified cross section of Delphi Technologies F2 family of Euro VI DFI

For the F2 family of products, Delphi Technologies has elected to use a solenoid stator to control the motion of a balanced 3-way hydraulic NCV. The electronically actuated Nozzle Control Valve (NCV) is the core of modern DFI technology, enabling precise and repeatable control of injection events. NCV performance needs to be balanced with leakage, both in terms 'dynamic' leakage associated directly with injection events, and 'static' leakage, and several possible valve architectures exist. A paper by Hardy et al (Hardy et al, 2012) summarises the design rationale for selection of a control valve for CV applications. While it features static leakage proportional to rail pressure, the 3-way valve concept was demonstrated to result in a lower total leakage than 2-way valves in CV applications (ibid). As such, it is a 3-way valve that was selected for the F2 family of Euro VI products and will be the only valve type referred to herein. The solenoid stator is a cost effective and durable means of driving the NCV and the pressure balanced 3-way NCV concept allows for a compact solenoid to be used for packaging benefits, while representing a significant cost reduction compared to indirect piezo actuators.

The NCV controls the pressure in a control volume above the nozzle needle, indirectly controlling its motion. The needle motion, and therefore injection characteristics, are characterised by flow through the inlet and drain orifices of the control volume. The Nozzle Assembly features a nozzle needle, a nozzle barrel, a nozzle body (not pictured) and a piston guide. The nozzle barrel and body contain the high-pressure reservoir integral to the injection rate shape. The nozzle needle features a return spring to control the rate of the end of the injection and a portion that is guided by the piston guide body.

3.2.3 DFI operation

Figure 11 provides an overview of the operation sequence of a DFI with a 3-way pressure balanced NCV acting indirectly on the nozzle needle, providing a stylised visualisation of the system. The High Pressure fuel source in HPCRS systems represents the common rail acting as a pressure vessel. Both the NCV pin and the nozzle needle have return springs acting on them to provide force balance and allow for the injectors' repsonse to be tuned through changes in preloading.

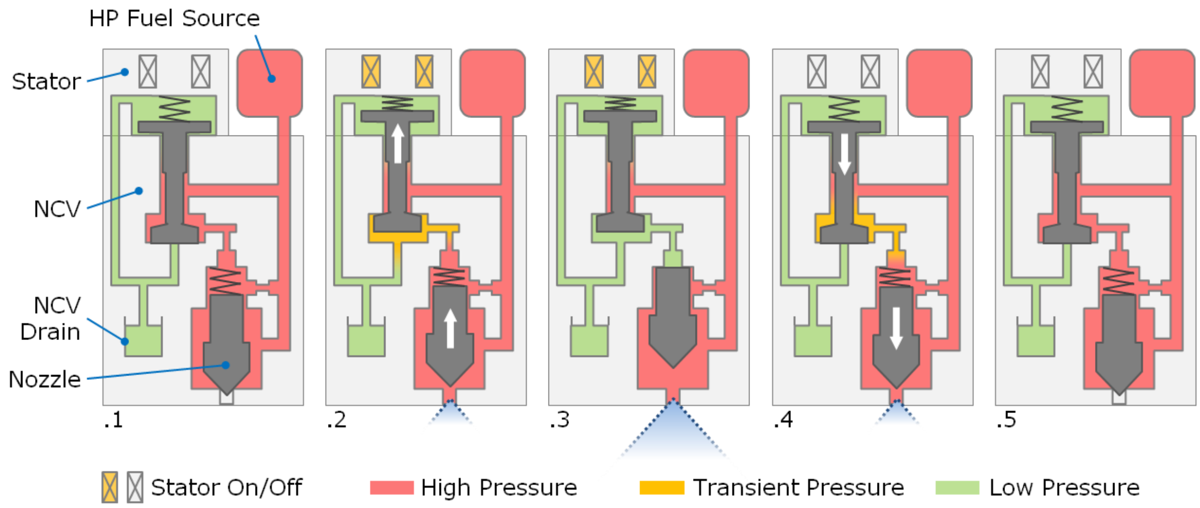


Figure 11: Operation sequence of DFI with 3-way pressure balance NCV

At rest, identified as stage 1 in Figure 11, the NCV pin is held at its bottom seat as a result of force balance, sealing the drain path preventing static HP fuel leakage. At this condition, the nozzle needle is also held hydraulically at its bottom seat, sealing the nozzle tip and preventing any injection from occuring.

At the start of injection, identified as stage 2 in Figure 11, the ECU energises the stator, creating electromagnetic flux. The resultant electromagnetic force acting on the NCV armature changes the valves force balance, causing the NCV pin to lift until it reaches its top seat. The NCV drain path is exposed, and results in the control volume being depressurised, changing the force balance on the nozzle needle such that it begins to lift of its bottom seat. High pressure fuel inside the nozzle assembly then beings to escape through the nozzle spray holes, starting injection.

Once the nozzle needle reaches its top seat, identified as stage 3 in Figure 11, full injection rate is achieved. The stator will remain energised by the ECU until the injector duration is sufficient to achieve the total fuelling demand. With both the NCV pin and needle on their respective top seats, the NCV drain is sealed from high pressures, reducing leakge.

When the required injection duration is achieved, the ECU will cease energising the stator, and the resultant decay in electromagnetic force will change the force balance on the NCV, causing it to travel back to its bottom seat, as identified as stage 4 in Figure 11. By sealing off the NCV drain path, the nozzle control chamber starts to repressurise, changing the pressure balance on the nozzle needle, causing it to start to return to its bottom seat.

Once the nozzle needle has returned to its bottom seat, injection will cease as indentified as stage 5 in Figure 11. With both the NCV bottom seat and the nozzle bottom seat sealed, high pressure fuel is sealed and no injection can occur. The pressure in the control chambers and nozzle volume will return to common rail pressure in preparation for the next injection event.

3.3 Legislative and customer drivers for Diesel FIS

Advances in Diesel FIS technology have enabled the continuation and significant development of the diesel Internal combustion Engine (ICE) paradigm, enabling modern engines to generate more useful work, while emitting significantly less harmful emissions. While the focus has historically been on increasingly stringent emissions legislations, several legislative and end-user focused market drivers influence the development of Diesel FIS. Figure 12 provides an overview of those market drivers, while Appendix I provides additional detail around each.

Figure 12: Market Drivers for CV FIS

The main driver for Diesel FIS has historically been to enable OEMs to meet increasingly stringent emissions standards, although there is evidence of a paradigm shift towards Fuel Economy and CO_2 emissions on the horizon. Euro VI and its equivalent emissions legislations also place further requirements on the OEMs to demonstrate emissions durability over extended durations, necessitating a FIS system that provides stable fuelling over life. Through the introduction of Euro VI and its enhanced NVH legislation, noise as a result of fuel pressurisation and combustion has become a more significant challenge to OEMs. To meet emission requirements in different applications, OEMs have often implemented different EATS strategies requiring different combustion strategies for maximum efficiency, placing differing injection rate shape demands on FIS systems. Finally, the costs associated with developing new engine platforms have led to OEMs being receptive to FIS that allow them to retain their existing engine architectures for subsequent emissions legislations. Diesel fuelled ICE are expected to remain part of the powertrain mix through to 2050, with a focus of efficiency improvements rather than further region-wide emissions legislation.

It can be suggested that the industry is in the process of a paradigm shift towards a focus on fuel consumption and Total Cost of Ownership (TCO), while maintaining low emissions, and improving reliability and weight. This suggestion is supported by the Automotive Council UK, who’s 2017 Roadmap for CV & Off-Highway Vehicle technology anticipates no further next stage ‘Euro VII’ emissions legislation, rather an increase in low and zero emissions zones (Automotive Council UK, 2017). The roadmap also predicts that ICE will remain part of the powertrain mix through to 2050, with further powertrain efficiency improvements, in part enabled by FIS technology, a key technological enabler.

3.4 The product characteristics employed by Delphi Technologies to meet the market drivers

The CV market has met Euro VI legislation through a combination of improved engine out emissions and advanced EATS. Two solutions are typically employed, either using increased Exhaust Gas Recirculation (EGR) rates to decrease engine out emissions or using increased levels of SCR (SCR), each placing different demands on the FIS. A visualisation of the typical system solution is shown in Figure 13, summarising the injection strategy and EATS typically employed in a Euro VI CV application.

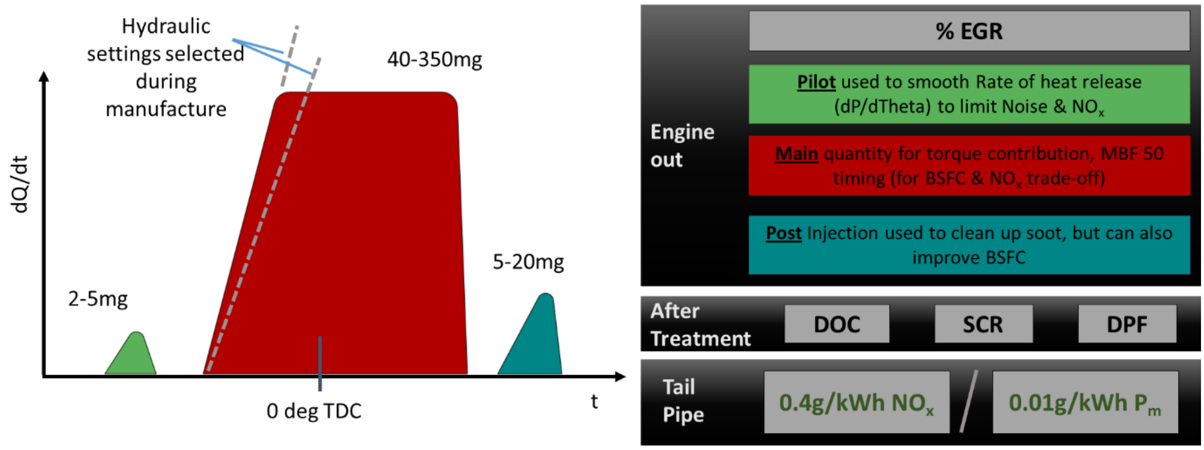


Figure 13: Typical System Solution for Euro VI

Delphi Technologies has developed the F2 family of products to enable its OEM customers to flexibly meet Euro VI legislation, while offering additional benefits in fuel consumption, noise, reliability, and packaging. A detailed overview of the technologies employed by Delphi Technologies in the F2 family of products can be found in the literature (Hardy et al, 2012) (Graham et al, 2014. Figure 14 shows the relationship between the market drivers discussed in §3.3, and the product characteristics associated with the F2 family of EU VI products developed by Delphi Technologies. The figure then shows the specific product properties that enable those product characteristics.

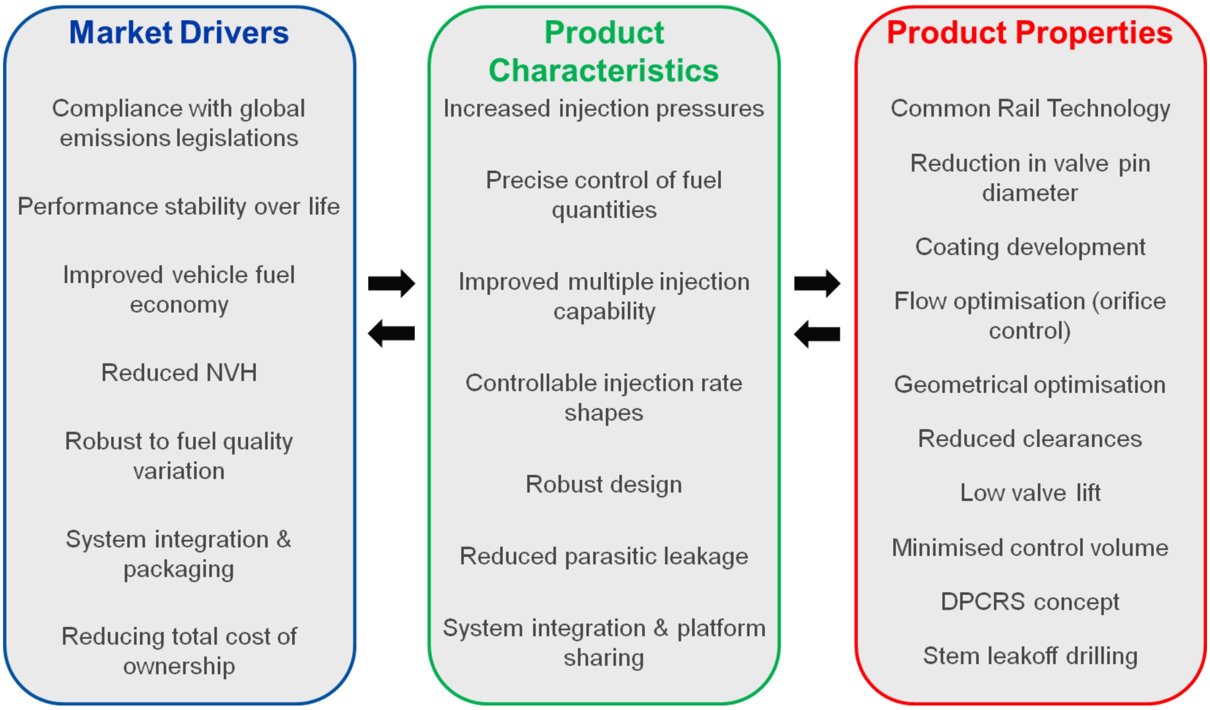


Figure 14: Relationship between Market Drivers, Product Characteristics, and Product Properties for CV FIS

The F2 family of products therefore differ in three fundamental ways to previous Delphi Technologies CV FIS systems. Firstly, maximum system pressures have increased to allow optimisation of emissions, fuel economy, and powertrain system cost, while common rail technology has been introduced. This results in increased pressure differentials across seats and increased hydraulic forces in the injector.

Secondly, the control of multiple injections has been improved with respect to quantity and timing, supporting advanced combustion strategies for optimisation of emission, combustion noise, and fuel economy, increasing the number of injection events for a given application and necessitating control of a wide range of injection quantities.

Finally, in order to reduce parasitic leakage to near zero the injector control valve has been miniaturised in conjunction with decreased design and manufacturing tolerances; this necessitates minimal wear to ensure performance stability over life.

3.5 Diesel fuel properties and their relationship with injector wear

The properties of diesel fuel, as both the working fluid and source of lubrication for mechanical interfaces, have a significant influence on the wear of fuel injectors. At the same time as technological advances have influenced the operation of FIS, the composition and characteristics of diesel fuel have changed over time, evident in the development of additives, the reduction in sulphur content, and the addition of biodiesel content. While the design of a DFI is the primary factor associated with its wear over life, the properties of the diesel fuel used in application are a significant contributory factor.

The properties of diesel fuel that influence diesel FIS durability can be divided into three distinct groups. The first group includes kinematic viscosity and lubricity, the primary characteristics that affect wear and durability. The second group comprises parameters other than viscosity and lubricity that effect FIS system performance over life, such as sulphur and Fatty Acid Methyl Ester (FAME) concentrations. The third group concerns any contaminations to the fuel, including hard particles and ingress of fuel or lubrication oil. An additional contribution factor is associated with degradation of the fuel associated with aging and recirculation. Furthermore, fuel properties can be shown to vary with temperature (Lacey et al, 2001). Table 2 provides a summary of the fuel properties relevant to FIS wear, while further details can be found in Appendix II.

Group & Fuel Property	Description	Effect on FIS wear
1 - Kinematic Viscosity	Resistance to shear forces	Poor hot weather performance; inability to reach high pressure; poor lubrication
1 - Lubricity	Resistance to wear of sliding interfaces	Sliding wear leading to loss of performance or seizures
2 - Initial Boiling Point	Temperature at which fuel evaporation begins at atmospheric pressure	Cavitation erosion leading to loss of performance or precipitants
2 - 95% Distillation Temperature	Temperature at which 95% of the fuel will have evaporated: indicator of non-volatile components	Fuel deposits
2 - FAME Content	Biodiesel content of the fuel	Fuel degradation leading to deposits
2 - Sulphur Content	Sulphur content of the fuel that results in SOx emissions	Corrosive wear to FIS
3 - Particle Contamination	Particles (soft & hard) within the fuel: indicator of cleanliness	Wear to FIS leading to loss of performance; fuel filter clogging
3 - Elemental Contamination	Concentrations of inorganic elements	Fuel deposits
3 - Water Content	Presence of water as dissolved, emulsified, or free-water	Wear to FIS leading to loss of performance; corrosion; microbial growth in tanks

Table 2: Fuel properties relevant to FIS wear

Using the fuel properties relevant to FIS wear, Delphi Technologies has divided the world markets for on-road diesel fuel into three groups: Category A markets are those with advanced quality, which meet the European EN590 standard for diesel; Category B markets are reflective of developing markets; and Category C markets have the worst fuel quality.

With respect to fuels for FIS validation on hydraulic test rigs with no combustion of the fuel, Delphi Technologies assess fuels based on both group 1 and group 2 fuel properties. The other fuel properties that influence FIS wear are either not considered as the effects are only visible on fired engines or are not typically considered as evaluation requires a specialized test with doped fuel. Using these criteria, fuels can be selected for validation testing to be representative of specific applications.

Given the potential for a wide range of fuel properties to influence injector performance over life, it is clearly a significant variable in NPD process. An understanding of the fuel used in applications, and typical market variations, is therefore a critical part of the PV process for any given application.

3.6 Potential implications for injector performance stability

The market has driven FIS technology to extremes in rail pressures and design clearances to balance fuel economy and engine-out emissions to suit a range of Euro VI strategies. The advanced combustion strategies used for Euro VI CV applications necessitate control of injection quantities and timing over an extreme range, requiring precise control of injector performance. Similarly, there is an increased requirement for injector stability over life to comply with emissions durability legislation. These factors combine to present a significant challenge to FIS manufacturers.

3.6.1 Increasing system operating pressure

Increasing injection pressures continue to be used to improve engine-out emissions and fuel consumption. By increasing the injection pressures, hydraulic forces in the injector are increased, increasing the loads associated with valve actuations and pressure dilation. At the same time, increased system pressures result in increased internal injector temperatures, increasing the propensity for fuel degradation and Internal Diesel Injector Deposits (IDID) formation.

3.6.2 Improved multiple injection capability; reduced parasitic leakage

The drive to increase, or at least maintain, hydraulic efficiency at increased pressures, results in FIS manufacturers utilising reduced clearances in design and manufacture. This is evident in the miniaturisation of control valve stem diameters and associated collar clearances utilised to ensure efficient operation at ultra-high rail pressures, while the fast-acting nature of the valve required for advanced combustion strategies results in a minimised valve lift. Through reducing design clearances, the systems have the potential to be more sensitive to changes associated with wear or IDID formation.

It can therefore be seen that in general the technological trajectory of FIS has led to a potential for decreased stability over life for Euro VI injectors, contrary to a key market driver. When combined with the requirements for precise control of a wide range of injection strategies, and the emissions durability legislation, unless the product design can demonstrate robust solutions to key failure modes, there is the potential for Euro VI injectors to experience warranty failures associated with that decreased stability over life. This is visualised in Figure 15.

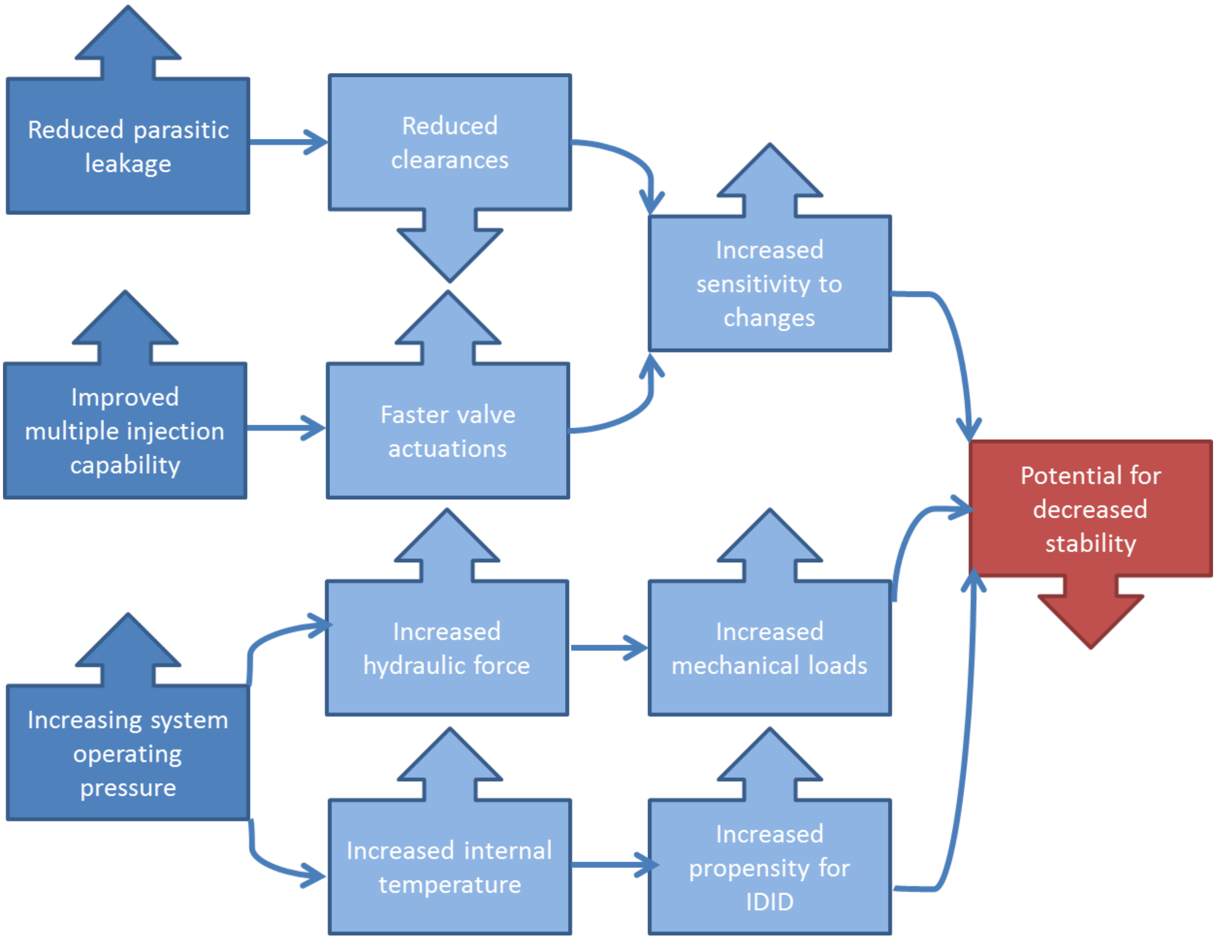


Figure 15: Potential influence of technical trajectory on FIS wear

3.7 Industry wide Fuel injector failure modes

This section will discuss the typical failure modes associated with diesel fuel injectors as reported across the industry. The three typical failure modes can be classes as deposits, fatigue failures, and valve wear, and specific focus will be placed on discussing valve wear, drawing additionally from analogous systems.

3.7.1 Fuel Deposits

Fuel deposits are a resultant of chemical reactions within the fuel or its additives, either resulting from oxidation in storage, contaminants in the fuel, or through pressurisation and the resultant increase in temperature associated with operation. The products or those reactions can either then be injected for combustion, recirculated through leakage or spilt fuel, or adhere to surfaces of the fuel injector when conditions are suitable, the latter of which are described as fuel deposits (Lacey et al, 2012). Fuel deposits are typically categorised as either being internal, or external, with regards to the injector. IDID are witnessed on the internal surfaces of the diesel injector, including nozzle needles and control valves. External deposits are typically associated with the nozzle but can also be exhibited on the capnut. While deposits on the injection nozzle, referred to by the industry as nozzle coking, have been witnessed for a number of years, IDID are a relatively modern phenomenon, with the potential for impacting injection performance over life, and therefore the emissions compliancy of an engine.

3.7.2 Fatigue failure

Owing to the high pressures associated with modern FIS, stresses associated with operation are such that fatigue failure is a possibility if design strength is insufficient. Changes in operation, such as the change to common rail systems where high pressures are present for longer, and start stop applications, which increase the number of low-cycle fatigue events a system will be subject to in application, also present different considerations with regards to fatigue. For the nozzle body, the effects of combustion are an additional consideration, with the possibility for embrittlement or corrosion to occur. Depending on the type and location of any fatigue failures, the resulting could be a loss of pressure resulting from either an internal or external fuel leak, or a failure that detrimentally influences other engine systems, including combustion. Fatigue strength of FIS products is therefore a significant design consideration, and the development and control of manufacturing processes to enable sustainable high strength products a key competency of FIS manufacturers.

3.7.3 Mechanical wear & material removal

As with any dynamic mechanical system, there is the potential for mechanical wear at component interfaces, or other forms of material removal, to influence the performance of a DFI. As previously discussed, the two sub-assemblies that influence the performance of an injector are the NCV assembly, and the nozzle assembly, both of which feature components travelling between different seats, and fuel flows. As such, both assemblies are susceptible to mechanical wear and other forms of material removal.

Classification of wear modes

Both the NCV and nozzle assemblies are valves that travel between two seats in operation and the interaction between the valves and their respective seats can result in both **impact wear**, from energy transfer on closing, and **sliding wear**, from relative displacement between the sealing surfaces resulting from elastic deformation under load.

While manufactures of Diesel FIS specify fuel cleanliness and filtration requirements to be warranted, it is always possible for there to be particulate contamination, or debris, within the fuel system. Particles with a low hardness are of low concern to injector components, but hard particles, with a hardness of an equivalent or higher level to the materials present in the injector, can result in **debris damage** and material removal upon impact with components which is of most concern on the valve seats of the NCV and nozzle assemblies.

As high-pressure fuel flows over valve seats and through the control orifices, and nozzle spray holes of an injector, it is possible for rapid changes in pressure to result in cavitation in the fuel, resulting in the formation of vapour bubbles on fixed surfaces. As those cavitation bubbles collapse under pressure, a shock wave forms that then transfers energy into the associated fixed surface, with the potential to result in material removal over time: this wear mechanism is known as **cavitation erosion**.

As a result of the significant pressure differentials associated with Euro VI FIS systems, fuel flows across valve seats and through the nozzle spray holes are increased. As the fuel flows across those pressure differentials, the energy transfer can result in **flow erosion** in the injector components as a result of hydraulic action, resulting in material removal.

Furthermore, it is possible for a combination of these failure modes to result in compound wear to the injector components. For example, hard particle debris damage can result in material removal or a valve seat, exposing relatively softer material under DLC layers, increasing the propensity for flow erosion to result in further wear.

Table 3 presents a summary of the different wear modes associated with injector valve seats, along with a brief introduction to the variables that can influence their severity. Wear to the valve seats can result in an increase in effective valve lift, a change in effective sealing diameters, inability to fully seal at high pressures, and/or changes in fluid dynamics across the valve seats. Such wear can influence the static and dynamic behaviour of the valves, changing the performance of the injector.

Wear Mechanism:	**Description:**	**Influenced by:**
Impact wear	Valve pin impact	Forces, mass, damping, seat alignment
Sliding wear	Elastic deformation of seat(s)	Forces, seat stiffness, seat alignment
Debris damage	Hard particle impact or trapping	Fuel cleanliness, filtration system efficiency
Erosion	Resulting from fuel flows or cavitation	Fuel properties, system design, seat profile design

Table 3: Wear modes of valve seats

As shown in §3.2 the NCV represents a form of poppet valve, with an axially loaded pin making a circumferential line contact over an orifice through differentially angled seats. Poppet valves are typically employed in a number of applications, including inlet and exhaust valves for internal combustion engines, medical devices such as replacement heart valves, and more widely in hydraulic power systems. Poppet valves are typically used in pressure control, and speed of response is a key characteristic (Bernad & Susan-Resiga, 2012). Poppet valves require resistance to wear and material removal resulting from operation (Blau & Hanft, 1993). Wear of this type of valve is often known as 'valve recession', with material removal at the seat resulting in an increased travel for correct valve operation.

Within the automotive domain, the wear of valve seats, predominantly inlet valves as exhaust valves are lubricated by combustion product, is a point of concern for engine performance, influencing both oil consumption and exhaust emissions (Lewis et al, 1999). Several potential wear modes have been identified, including impact wear, sliding wear, abrasion, adhesion, and corrosion. Inlet valves are subjected to both an initial closing load associated with the cam design, and a subsequent, higher load associated with combustion while high temperatures associated with combustion can also result in distortion of the cylinder head valve seat, with the potential to result in an increase in both impact and sliding wear.

Through structured empirical investigation, Lewis et al (1999) determined that the prevailing wear mechanisms were related to both design and usage variables. Furthermore, by isolating impact and sliding wear, the investigation demonstrated that both wear mechanisms result in material removal, but the magnitude of the total wear is greatest when the two mechanisms are combined, resulting in compound impact wear. Lewis suggested that "the most powerful way of reducing wear is to keep impact velocities as low as possible".

Lewis and Dwyer-Joyce (2002) followed earlier work with a combination of empirical and analytical investigation to further investigate the relationship of design solutions and usage conditions on inlet valve wear. Valve seat wear was modelled in terms of sliding and impact wear, using Archards equation (Archard, 1953) to model the abrasive wear associated with sliding wear (Suh & Sridharan, 1975), and an equation used to describe the impact wear of poppet valves presented by Fricke and Allan (1993) for poppet valves in mining equipment. The valve recession model presented by Lewis & Dwyer-Joyce was demonstrated to provide a quantitative predictor of valve seat recession, with components of both forms of wear, for a given valve seat material pairing, and under certain operating conditions.

Furthermore, the empirical testing demonstrated that valve recession increased by ~3.5x when no lubrication was present (Lewis & Dwyer-Joyce, 2002). Lewis & Dywer-Joyce went on to suggest that the form of lubrication present is boundary lubrication, owing to the relatively slow sliding velocities and high axial loads present, with the resultant film thickness too thin to provide complete separation of the surfaces.

Lewis (Lewis, 2007) suggests that impact wear had not been studied extensively, resulting in a relative lack of information on root causes and wear mechanisms, before developing a modelling technique to predict compound impact wear, correlated with empirical results.

Using the model first proposed in the work of Lewis & Dwyer-Joyce (Lewis & Dwyer-Joyce, 2002), Lewis further validated the model for compound impact wear using empirical data from the literature (Lewis, 2007). The model was shown to demonstrate good correlation with the test data, except for low cycles, which was suggested to be a result of the model not predicting any plastic deformation, only material removal.

3.8 Summary

This Chapter has provided background information through an outline of the market drivers that influence the development of FIS and introduced the product characteristics and properties that have been identified as technical responses to those drivers. Miniaturisation of the control valve and associated hydraulic volumes, alongside an increase in system operating pressure, has been presented as a typical technological solution to multiple market drivers.

The influence of fuel properties has also been discussed with reference to the robustness of FIS, with potential to influence a number of different possible failure modes. The potential implications on the robustness of FIS was also introduced, alongside an introduction to product failure modes evident in the wider industry.

The next Chapter will provide the background relating to the main academic domains associated with this research, and the specific research methods employed in this research.

Chapter 4 Review of the Literature

4.1 Introduction

This Chapter presents the background theory, research, and information associated with the three main academic domains associated with this research: Expert Judgement and Elicitation; Experimental Design; and Reliability Engineering. An overview of the background associated with the application and management of knowledge in the new product development process will be presented in §4.2, with specific focus around the role of expert judgement and how it can be elicited. §4.3 will then discuss System Dynamics models within the context of NPD.

The background theory associated with the Experimental Design methodology employed as part of this thesis will also be presented in §4.4. The methodologies available will be discussed with regards to the considerations and limitations associated with their design, usage, and analysis, drawing from academic and industrial applications in the literature.

Finally, the background theory associated with the accelerated life testing methodology employed by the industry to demonstrate the validity of new products with regards to both functional and reliability specifications will be presented in §4.5. The challenges and limitations associated with the methodology will be discussed, alongside the application of engineering judgement in the presence of uncertainty in the new product development process. In doing so, it will be shown that only through appropriate knowledge of a product, its application, and its failure modes, it is then possible to design controlled laboratory experiments such as to demonstrate product reliability in a resource sensitive manner, and that those experiments in turn generate significant new knowledge of a product.

4.2 Knowledge Management & Expert Elicitation

Knowledge is of key strategic importance for firms engaged in developing new engineering products; by enhancing a firm's core competencies, its ability to deliver new, improved products to market with lower resources required is increased. In the context of engineering systems, knowledge can be defined as being the implicit or explicit understanding of the 'why', that provides context and significance to objective data, gained through practical experience or through education.

New Product Development, and PV in particular, require numerous decisions to be made, often under conditions of uncertainty, based on the available data and the organisation existing knowledge. These decisions can 'lock in' an organisation in to a design or testing strategy, committing significant resources to its completion. As will presented in §4.5, a key element of the PV process is the design of the Accelerated Tests that will allow the organisation to identify and apply corrective actions to failure modes representative of the products end usage. Ultimately, the results of the PV process are used to demonstrate the product's reliability, a key metric of the product development process, and an indicator of future warrantee costs. The design of such tests is typically conducted using the results from limited testing of pre-production samples, augmented with prior engineering knowledge (Nelson, 1990). Those test designs then commit the organisation to providing the time, samples, engineers, and other resources required to complete the PV programme. As such, efficient utilisation of organisational knowledge plays a key part in the performance of the NPD programme, both in terms of a successful end products, but also in the resources required to take a concept to market (Erto & Giorgio, 2002).

When dealing with uncertainty in NPD, 'engineering judgement' is a phrase commonly used to describe the application of reasoning to inform a decision when objective data is unavailable or insufficient. More generally, Expert judgement facilitates the drawing of rational conclusions from the available data under conditions of uncertainty. In doing so, knowledge will have to transition across the boundaries (Carlile, 2002) that exist within a business, with a potential for loss, and mistranslation in the process. Many experts may be consulted for one problem, but how is the anecdotal knowledge provided by those experts assessed for accuracy and reliability, and how is it best captured for subsequent re-use and combination?

Expert Elicitation describes the process by which appropriate expert judgement is identified, gathered, and prepared as an input to the decision-making process. Effective and timely expert elicitation, and the quantification of uncertainty, is of key importance in product engineering, as over the course of an NPD programme, countless product design decisions are made with engineering judgement as a significant consideration. The available data is codified in models, calculations, and reports, but the engineering knowledge that helped informed those decisions is infrequently captured. When new data is made available, it can therefore be difficult to reassess the integrity of those decisions, and the business can fail to learn.

4.2.1 Knowledge & Knowledge Management

Given a level technological & financial playing field, it is knowledge that separates two competing engineering firms (Sainter et al, 2000). While technological advances have allowed a step increase in the amount of data generated through the design and evaluation of an engineering system, it is often knowledge that is relied upon in decision making (Venkatubramanian, 2009). However, knowledge can act as both a source, and a barrier for innovation in NPD, so requires consideration and management (Carlile, 2002). The challenge is therefore how to best manage knowledge in NPD programmes, and how to leverage that knowledge to enable a market advantage.

Ragab & Arisha presented a comprehensive review of the literature associated with Knowledge and Knowledge Management (Ragab & Arisha, 2013). Their review presented the differing definitions of knowledge, suggesting the hierarchical definition, where data becomes information, which in turn becomes knowledge, is the most accepted. Their review also discusses the classifications of knowledge presented in literature, presenting the dichotomy of tacit and explicit knowledge as the most commonly used. Ragab and Arisha also describe how knowledge has been viewed as the current of the modern economy, and the key to creating and sustaining a competitive advantage in the market.

Knowledge Management (KM) has been widely discussed in the literature since the 1990s. The discipline has focused on transferring tacit knowledge to explicit knowledge, described using the knowledge spiral of Socialisation-Articulation-Combination-Internalisation (SACI), through which an individual's tacit knowledge can be captured in an organisation's products & processes, and in turn increasing the organisations knowledge capacity (Nonaka & Takeuchi, 1995). In their review, Ragab & Arisha provide an overview of the differing definitions of KM discussed in literature (Ragab & Arisha, 2013).

Numerous studies have investigated the positive effects, both qualitative and quantitative, of KM on the performance of NPD programmes in industry (Sherman et al, 2005) (Vaccaro et al, 2010) (Chang et al, 2005) (Liu et al, 2005). Furthermore, the process of engineering of complex products is more susceptible to changes in the project team, the organisation, and to the project's context, and as such, the possible contribution of KM to risk reduction is of increased significance (Cooper, 2003).

In their review of the literature, Ragab & Arisha (Ragab & Arisha, 2013) presented the discussion of the managerial and social issues concerning KM. Their review concluded that the main problem for organisations is the tendency for employees to 'hoard' their knowledge and the review went on to summarise how organisations can motivate knowledge sharing through KM interventions.

4.2.2 Expert Judgement Elicitation

Appropriate use of expert judgement can provide decision makers with knowledge supplementary to objective data and has been described in numerous contexts in literature. Cooke (Cooke, 2013) suggests that expert judgement helps an organisation more towards a rational consensus from evidence in the event of uncertainty. Similarly, it has been described as essential in the event of incomplete data, in unique contexts, or when extrapolation outside of an existing body of knowledge is required (Burgman et al, 2011). In the context of NPD programmes, expert judgement is suggested as a means for assessing uncertainties that emerge during the design process (Bedford et al, 2006), and for structuring and parameterising a useful model in knowledge intensive processes (Ford & Sterman, 1998). For this research, elicitation of expert judgement is considered to facilitate informed decisions when uncertainty is abound.

Expert judgement is elicited more formally through the product design process, with the two most widespread examples being QFD, and FMEA. In both methods, the judgements of a cross section of experts are called upon to augment codified knowledge, and to provide a substitute in the case of uncertainty. However, when it comes to the design of an appropriate empirical testing programme, there is little evidence that such methods are adopted in such a widespread manner. Empirical testing programmes associated with product validation, particularly reliability demonstration programmes, can be highly resource intensive, and, if inappropriate design leads to the discovery of new knowledge late in the product development cycle, it can be expensive to act upon. In order to best design the Accelerated Tests associated with product validation, knowledge of the failure mode(s) of the products is required (Nelson, 1990), and if uncertainty exists around those failure mode(s), expert judgement will be relied upon.

Expert Elicitation describes the process(es) through which expert judgement is elicited, such that is can be incorporated into a decision-making process. Expert Elicitation can occur either formally, or informally, but literature concurs that it represents a worthwhile tool in the decision-making process, and expert judgement should be elicited frequently, regardless of formality (Kadane & Wolfson, 1998). An example of an informal application can be seen in the typical experimental process, where a problem is articulated, experimental boundaries are identified, and appropriate variables are selected; in doing so, the experimenter(s) draw upon their own expert judgement, and that of their peers, in making those informed decisions (Ford & Sterman, 1998). In such an informal application, there are two roles: that of the decision maker who has the problem, and that of the experts that have the appropriate subject matter expertise (Bedford et al, 2006). In more formal applications, Expert Elicitation typically incorporates a third role, with the addition of an analyst responsible for identification of the appropriate experts, and for facilitating the elicitation process (ibid).

It has been suggested that there are no fixed criteria for what constitutes an expert (Goossens et al, 2008), but that "it is reasonable to hope that they [the experts] will have though harder, and over a longer period of time, about the subject matter at hand than others have" (Kadane & Wolfson, 1998). Others have suggested that experts can be identified through a combination of their qualifications, track records, and relevant experiences (Burgman et al, 2011). As such, expertise in a subject may not necessarily correspond with seniority within a hierarchical organisation or peer group.

Regardless of the expert elicitation method employed, and the panel of experts selected, elicitation of expert judgement should be undertaken with the following principles in mind: ability to hold up to scrunity; fairness; neutrality; & performance control (Goossens et al, 2008).

Expert Elicitation is subject to the possible influence of different forms of bias, both in the expert panel, and in the analyst alike. Three forms of possible bias in expert panels have been identified: motivational bias in the form of vested interests; anchoring bias where inappropriate prior data is transformed to suit the problem case; and availability bias where emphasis is placed on the most memorable, if not pertinent, historic data (Bedford et al, 2006). The analyst can introduce bias into the expert panel through the sample data provided, particularly in the case of quantitative studies (Kadane & Wolfson, 1998).

Further studies have attempted to assess the reliability of the judgement provided by individual experts using paired comparisons (Goosens et al, 2008), but cross validation of experts has proven difficult to achieve (Cooke, 2013).

4.2.3 The Delphi Method

The Delphi Method is an established Expert Elicitation method which can trace its roots to Cold War-era US military contract work (Skulmoski & Hartman, 2007). Okoli & Pawlowski (Okoli & Pawlowski, 2004) describe the Delphi Method as a means for best structuring communication such as to allow a group to deal with a complex problem. The four key characteristics of the Delphi Method are: the anonymity offered to the panel; the iteration of opinion it encourages; its feature of controlled feedback through, and the aggregation of a groups response it enables (Skulmoski & Hartman, 2007). The Delphi Method has been identified as suitable for forecasting where knowledge is imperfect, there are no correct answers or hard facts, and group consensus is considered an acceptable choice (Donohue & Needham, 2009). It can prove suitable when subjective group consensus has more value to the decision maker(s) than incomplete objective data, or when personal contact with the experts is not possible or desirable (Bolger & Wright, 2011). The anonymity afforded by the method reduces the tendency for more junior experts to align their views with their senior colleagues and can lower the probability of experts going off on a tangential line of inquiry (Geist, 2010).

The Delphi Method has been presented as a proven means of identifying or quantifying variables, generalising theories, and understanding causal relations (Skulmoski & Hartman, 2007). Successful implementations have resulted in theory generation and have been shown to derive new insight into the subject rather than just being a means of data collection (Day & Bobeva, 2005). It offers a more rigorous methodology for eliciting expert judgement than a traditional survey, offering flexibility in design, and leveraging an aggregated panel response to answer complex questions appropriately

(Skulmoski & Hartman, 2007). A special case of the Delphi Method that has experienced widespread use is the 'Ranking' type Delphi study, were the analyst seeks to develop a group consensus on the relative importance of issues (Okoli & Pawlowski, 2004).

It is often mistakenly stated that the Delphi Method has the singular aim of reaching consensus on a subject, but instead the aim should be a stability in response, where, for example, a bipolar state of opinion amongst the expert panel is a valid and important result (Linstone & Turoff, 2011). By reaching such a stability, the Delphi Method can alert the decision makers, and the expert panel alike, to the complexity of a problem, and challenge their assumptions (ibid). Therefore, the value in the method is in the ideas it generates, both when consensus is reached, and when informed disagreement between experts is witnessed (Gordon & Pease, 2006).

4.2.4 Criticisms of the Delphi Method

In the 50 years that the Delphi Method has been used in the public domain, numerous criticisms have been levied at the methodology. It has been suggested that one of the key characteristics of the method, its avoidance of face-to-face interaction, is also one of its most significant weaknesses, allowing experts to avoid direct challenges to their judgement (Patton, 1997).

Additional methodological weaknesses have been presented, including the time and effect required to complete a multi-round study, the ease with which the analysts bias can influence the study, and the difficulty in confirming the accuracy and validity of the results (Landeta, 2006). Additionally, literature has shown that applications lacking in rigour can lead to unfavourable results, through inappropriate question design, poor timing, selection and management of the expert panel, and insufficient analysis of the results (Skulmoski & Hartman, 2007) (Geist, 2010) (Landeta, 2006). Furthermore, the premise of the study assumes that the individual panel members are open to having their opinions challenged and changed by their peers (Donohoe & Needham 2009).

4.2.5 Best Practice for Successful Delphi Studies

The literature has demonstrated that the initial contact with the expert panel, in terms of both the quality of the material provided and the engagement levels achieved, highly influences the potential overall effectiveness of a study. A formal invitation, detailing the research motives, participation requirements for the study, and a description of the expert selection criteria has been shown to be effective in engaging with the expert panel (Donohoe & Needham, 2009). Given the potential for availability constraints for typical experts, clarity of the time requirements for study participation is suggested to be provided upfront (Skulmoski & Hartman ,2007). Okoli & Pawlowski (Okoli & Pawlowski, 2004)recommended that the first round of the study should be possible to complete within

30 minutes while providing an upfront expectation of the total time requirements associated with participation.

Once initial panel engagement has been achieved, several strategies have been identified to improve panel response rate by improving panel retention, including ensuring a quick turnaround between rounds to maintain momentum, and limiting the duration of the panel to limit fatigue (Gill et al, 2013). Furthermore, it has been suggested that repeated communication with the panel members being described as 'experts' is a means for best ensuring continued engagement (Gordon & Pease, 2006).

Selection of the panel size is another important consideration affecting both the potential results of the Delphi Method and the efficacy of its application. Large panels have the potential advantage of aggregating the judgement of a wider number of experts but can be difficult to manage and can often experience high attrition rates (Gill et al, 2013). Conversely, small panels, while easier to manage, can potentially result in limited generalizability of the results, with bias in smaller number of experts influencing the aggregated response (ibid). The minimum panel size recommended by Donohue & Needham (Donohue & Needham 2009) is between 7 and 15, although homogeneous panels can be smaller, while Okoli & Pawlowski (Okoli & Pawlowski, 2004) recommend 10 to 18 members.

4.2.6 The Nominal Group Technique as an Alternative to the Delphi Method

The Nominal Group Technique (NGT) is a method for EE that has some similarities with the Delphi Method, but presents different considerations for its usage. NGT is a group-based process for sharing a problem, proposing and developing a number of possible solutions, and then voting to prioritise or rank those solutions (Delbecq et al, 1975). Similar to the Delphi Method, the NGT provides a means to avoid the problems typical of group interactions, by providing an equal platform for all group members to contribute, and in avoiding conflict through structured facilitation.

However, compared to the Delphi Method, the NGT has the limitation requiring physical attendance of a group of experts to workshops, which would be multiple if iteration was desired. The Delphi method also offers anonymity that is not present in the NGT, with the potential for improved results dependant on the expert panel members and organisational norms.

4.3 System Dynamics Modelling

System Thinking is a methodology that integrates several disciplines, leveraging logic in the understanding of patterns and interactions in complex problems, providing a framework for understanding the complexity and dynamicity of systems (Cavana & Mares, 2004). System models are a tool typically employed in the practice of System Thinking to help understand systems, and their dynamic behaviours. Haraldsson (Haraldsson, 2000) suggests that mental models tend to be constrained to static simplifications of a system, and it is the dynamic complexity that one seeks to understand when engaged in complex problems.

In his review of the development and application of System Dynamics models, Lane (Lane, 2008) describes the System Dynamics domain to be diverse, but suggests that in most applications its purpose is to support decision making processes and organisational learning, going on to say that while it has links with the ‘soft’ systems methodology, System Dynamics itself facilitates an integration of aspects from both the ‘soft’ and ‘hard’ paradigms. Lane also suggested that the purpose of their use is to help understand system behaviour, such as to explore suitable interventions, and that models ultimately serve to transform uncodified knowledge into mathematical models of a system (ibid).

The literature provides a comprehensive overview of the elements of Systems Dynamics models in general terms, including variable types, flows, and stocks (Lane, 2008) (Haraldsson, 2000). Lane (Lane, 2008) suggests that the identification of the polarity of feedback loops within a system is a crucial step towards understanding the dynamic behaviour if a system, identifying whether a perturbation results in an response by the system, or whether it remains at equilibrium.

Morecraft (Morecroft, 1982) suggests that a Systems Dynamic modelling tool has the requirement of being an efficient means for organising the qualitative data associated with mental models of dynamic systems, such that an overview of the system structure & response can be portrayed. That efficiency is manifested through an accurate portrayal of the processes of the system, and through expressing the system in such a way that is directly comparable to existing mental models (ibid).

4.3.1 Causal Loop Diagrams

Causal Loop Diagrams (CLD) are one of the predominant forms of Systems Dynamics models in literature (Morecroft, 1982), owing their popularity in part to their utility in making system dynamics accessible to a wider audience (Richardson, 1986). Haraldsson (Haraldsson, 2000) provides a comprehensive explanation of the concepts associated with CLDs while providing a good description of how they fit in with mental models of system dynamics.

They provide a structured format for articulating an understanding of the dynamic nature of a system (Kim, 1992). CLDs need not be fully formed such as to contain all variables typically required of a model, allowing a simpler representation with more focus on the structure of feedback loops (Lane, 2008). While not being part of Forrester's original schema, CLDs have emerged from the development of the System Dynamics domain to take a significant role in the exploration and communication of systems (Morecroft, 1982).

4.3.2 Benefits of CLD

The literature describes the advantages of CLDs as being associated with their notation, and the low resources required for their implementation. Lane (Lane, 2008) proposed that the simple language used in the models allowed for an organic introduction of the models to discussions, while the reduced complexity of the model makes them useful for communicating the system response to untrained audiences, with loops clearly communicated. Furthermore, their low resource requirements make them suitable for use as a quick tool for exploring a system, with multiple iterations possible in quick succession, requiring only paper or a whiteboard for implementation (ibis). Lane also suggested that by limiting detail, CLDs place increased focus on feedback and decision points within a system and that the explicit communication of the location of feedback loops can also be of benefit in clearly identifying significant feedback structures that may otherwise be missed (ibid).

Yearworth & White (2013) concluded that the iteration between the grounded data, CLD, and ultimately, formal mathematical models, can result in further insight into the system, generating additional research questions.

4.3.3 Criticisms of Causal Loop Methodology

When reviewing the modelling techniques emerging in the Systems Dynamics methodology, Morecroft (1982) presented several perceived weaknesses associated with CLDs. Those weaknesses included the gap between people's mental models and the loop structure of CLDs, suggesting that people's mental models of sociotechnical systems would not tend to extend to how its components interact. Among the other weaknesses presented by Morecroft was the lack of definition in representing different variable types within the system.

A paper by Richardson (1986) criticised CLDs for being open to inconsistencies in notation and interpretation of the +/- polarity of nodal links, both additive and proportional, and went on to propose an alternative notation that has not been widely accepted. Richardson concluded at the time that CLDs were a model format best placed for communicating with untrained audiences. In a later paper, Richardson (Richardson, 1997) went on to criticise the alternate same/opposite (s/o) notation that was popular in practitioners at the time, concluding that "[the use of s/o] is logically flawed and must be ruled out”, suggesting the existing +/- notation was more suitable for the deep thinking associated with complex systems.

A more recent paper by Lane (2008) argues that the fundamental weaknesses of CLDs had yet to be addressed, and that the resulting models were weak in representing influences and could be inadequate at describing complex system behaviour. Those weaknesses were attributed as CLDs lack of precision in focusing on feedback, and the simplistic notation lacking a differentiation between stock and flow variables. Lane also suggested that while the suppression of fine detail in the model can make it attractive to non-trained audiences, it can result in system behaviour having to be inferred.

4.3.4 Best practice for applications of CLD

The literature presents a number of best practices associated with implementation of CLD. Kim (1992) presented a series of guidelines for the application of CLD, suggesting firstly that the selection of the system and the specific thematic ‘question’ that is to be explored is of upmost importance, and not all such questions are equal, and once they have been identified, a time horizon appropriate to the issue should be selected. Kim goes on to suggest that selection of an appropriate system boundary is another key consideration, suggesting that the expected system response associated with different variables is assessed for significance before addition to the model. When considering the level of detail associated with the model, Kim suggests that the practitioner should consider the thematic question and time horizon, as the behaviour of some variables may be so slow compared to the response of the system as to be suitable for consideration as constant.

Yearworth & White (2013) utilised binary matrices to define CLD based on the qualitative data defined through semi-structured interviews, before suggesting that most organisations already have an abundance of such data available, and such methods could be readily applied. They also suggested that once sufficient qualitative data had been coded into causal relationships, the significance of each such link should be examined, identifying its meaning and polarity of causality, with variables & links being of little-to-no significance to the system response being removed from the model as auxiliary data.

The literature also provides best practices for naming of the nodal concepts in a CLD, with Cavana & Mares (2004) recommending that concepts be operationalised as variables, while Kim suggested verbs and action phrases should be avoided as variable names to avoid confusion with the polarity of links, and the positive sense of variables should be chosen for the same reason (Kim, 1992). Kim went on to suggest that variables that can vary over time should be selected.

If a link between two nodes in a CLD requires additional explanation for clarity, the practitioner should revise the variable names, or insert an intermediate term such as to improve the readability of the model (Kim, 1992). The CLD should be simplified such as to be best used for understanding the system in order to identify suitable interventions, enabling the 'closed loop' thinking associated with the systems approach (Cavana & Mares, 2004).

4.4 Experimental Design Methodology

Experimentation is a key element of the scientific method, performed by investigators in all fields of study, in order to increase knowledge about a process or system (Montgomery, 2009). An experiment can be defined as:

> *"A test, or series of tests, in which purposeful changes are made to the input variables of a process or system so that we may observe and identify the reasons for changes that may be observed in the output response" (Montgomery, 2009)*

While some such responses can be explained through physical or chemical, mechanistic models, most responses require observation and experimentation in order to be characterised fully in empirical models (Montgomery, 2009). Experimentation has also been described as a knowledge creation process, leading to the development of a product, system, or organisation (Thomke, 2003), and the objective is therefore learning, and acts as a means for reducing uncertainty (ibid).

Product testing is a descriptor typically used to describe experimentation in an industrial context. Through improving the knowledge of a product in order to reduce any associated uncertainties, experimentation is a critical part of the NPD process, and is evident in each of the group problem solving methods that will be presented in §4.5.1. Product testing is the physical manifestation of experimentation in the NPD process, but is also evidenced virtually in the Computer Aided Engineering (CAE) tools and methods employed in the Design and Analysis of products. By enabling systematic verification of robust design solutions, experimentation is at the core of every company's capability to innovate, allowing them to create and refine their products (Thomke and Bell, 2001). However, it is suggested that experimentation in most NPD programmes is not done in an optimal manner (O'Conner, 2001) (Shabi and Reich, 2012), with the lack of a consistent methodology for experimentation in NPD cited as the main cause (O'Conner, 2001).

The perceived sub-optimality associated with most industrial experimentation is reflected in the two experimental design approaches most frequently employed in industry (Tanco et al, 2008). The first of such approaches is the 'Best Guess' approach, where the experiment will start with all input variables at the existing baseline condition. After each experiment, a combination of observation and expert judgement is then employed to change one variable. This best-guess approach may continue for prolonged durations, with neither a guarantee of improving the product or process, nor that the optimum solution has been reached if any such improvement is observed. This approach can be seen to be inefficient with respect to resource requirements, while providing no guarantee of success (Montgomery, 2009). Figure 16 visualised the best-guess approach to an experiment with 3 variables.

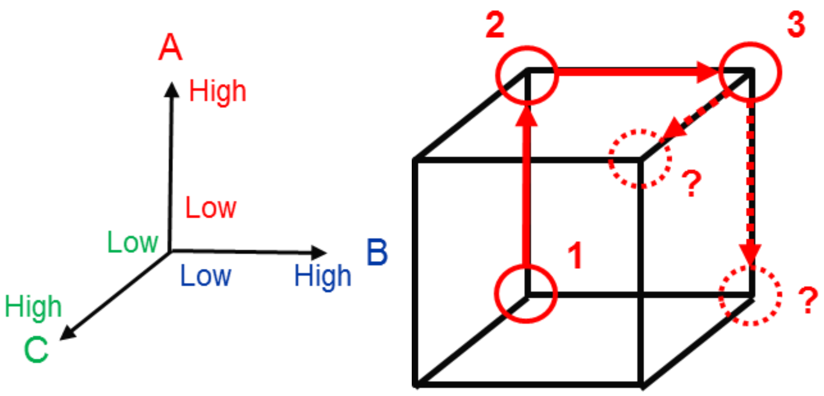


Figure 16: Best-Guess approach to an experiment with 3 variables

The second approach used extensively by practitioners is the 'One Factor at a Time' (OFAT) approach, in which, the experimenter varies each factor, one at a time, from a baseline condition, before then demonstrating the effect of each factor on the response variable(s) (Montgomery, 2009). The OFAT approach therefore provides a more structured approached to experimental design than the Best Guess approach (ibid). However, it does not account for interactions between factors, which are common, and can result in erroneous conclusions (ibid). In a similar manner to the Best Guess approach, the OFAT approach requires significant resources to obtain only a limited amount of knowledge about the product or process (Antony, 1998) and is considered an unreliable method for obtaining valid conclusions about design or process optimums (Logothetis, 1994). Figure 17 visualised the OFAT approach to an experiment with 3 variables.

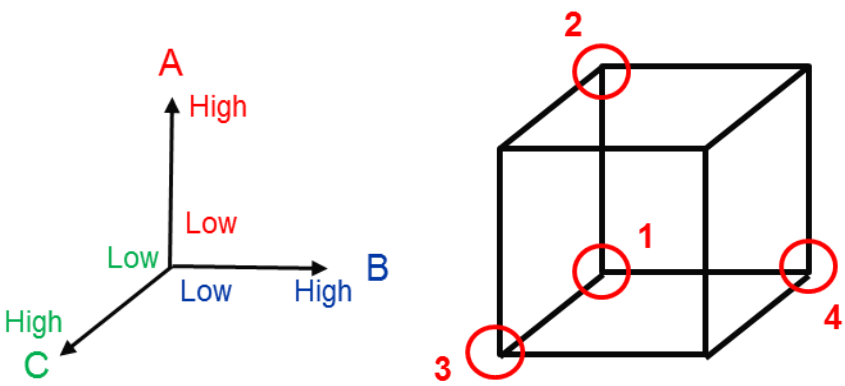


Figure 17: OFAT approach to an experiment with 3 variables

Both the Best Guess approach and the OFAT approach require significant expert judgement in the selection of appropriate test factors (Montgomery, 2009), and have been suggested as relying on guesswork, luck, and intuition for success (Hockman and Jenkins, 1994. Both approaches are always less efficient that alternative methods based on statistical approaches (Montgomery, 2009).

The Factorial approach, and the underlying statistical methods, is considered the 'correct' approach for experimental design (Montgomery, 2009). In the Factorial method, factors are varied both individually, and in combination such as to assess the contribution of not only each factor, but also their interactions (ibid). By allowing the experimenter to assess the effects of individual factors and their interactions, factorial experimental designs make the most efficient use of the experimental data (ibid).

4.4.1 Design of Experiments Methodology

Design of Experiments (DoE) methodology, also known as Experimental Design methodology (ED), presents a systematic and structured approach to experimentation (Antony, 1998). Factorial Designs form the basis of the DoE Methodology, representing a range of techniques used to understand the effects of different variables in the development of products and processes (Booker et al, 2001).

DoE methodology describes to the process of planning an experiment such that appropriate data will be collected for analysis by statistical methods, resulting in valid and objective conclusions (Montgomery, 2009). DoE is a considered a powerful technique for determining the optimal factor selection for a product or process, achieving improved performance and reduced sensitivity to uncontrolled factors (Antony & Antony, 2001). Through turning designed experiments into empirical system models, DoE is a form of a Knowledge Transformation (Montgomery, 2009).

There are 3 basic principles that govern the DoE Methodology: Randomisation, Replication, & blocking (Montgomery, 2009). Those three principles allow the practitioner to reduce, and quantify experimental error, while allowing a statistical comparison of the effects associated with the design variables to be made.

4.4.2 Factorial & Fractional Factorial Designs

Of the many experimental designs that study the effects of two or more factors, Factorial Designs are the most efficient as in any given trial or replication, all possible treatment combinations are tested while allowing for calculation of experimental error (Montgomery, 2009). In an OFAT experimental design, 2 factors at two 2 levels would require 6 tests to be completed in order to assess the effect of each factor, and to account for experimental error, while a similar factorial design would require just 4 tests (ibid).

A Factorial design of **k** factors at **n** levels each, is described as a n^k design. This simplest factorial design is therefore a 2^2 design, where two factors, A and B, have each associated two levels, nominally high and low (Montgomery, 2009). The standard nomenclature describes the effects associated with each factor, and their interaction, with capital Latin letters, such that in the case of a 2^2, there are effects associated with A, B, and AB. Figure 18 then provides a visualisation of a 2^3 design, with three factors, each at two levels.

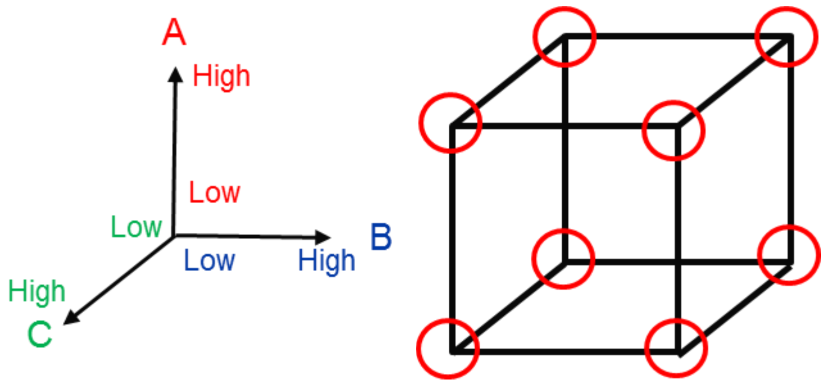


Figure 18: 2^3 Factorial experimental design

However, for even a two-level, 2^k, design, as the number of factors **k** increases, the number of treatment combinations, and therefore experimental runs required, increases exponentially. For example, a Factorial design to assess the effect of 10 factors and their interactions would require 1024 treatment combinations (2^{10}). Such an experiment would typically be unfeasible from both a time and resource perspective (Montgomery, 2009). One solution available to experimenters, is to use Fractional Factorial Designs (FFD), where only a subset of the runs of a Factorial design are used, as shown in Figure 19 for the case of a 2^{3-1} FFD (ibid). Fractional Factorials are used extensively in industrial research and development, lowering resource requirements, but introducing confounding of main effects with interactions.

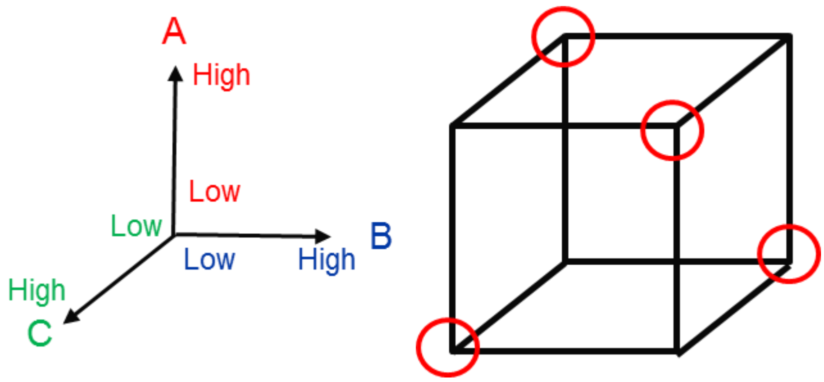


Figure 19: 2^{3-1} Fractional Factorial experimental design

When considering the common 2^k design, one possible limitation concerns linearity of effects: with only 2 levels for each factor, responses are assumed to be linear. This assumption can be tested through the inclusion of centre points, or pseudo centre points in the case of discrete variable levels, examples of which is shown in Figure 20 for a 2^2 with continuous variables (left), and a 2^{3-1} with one discrete variable (right).

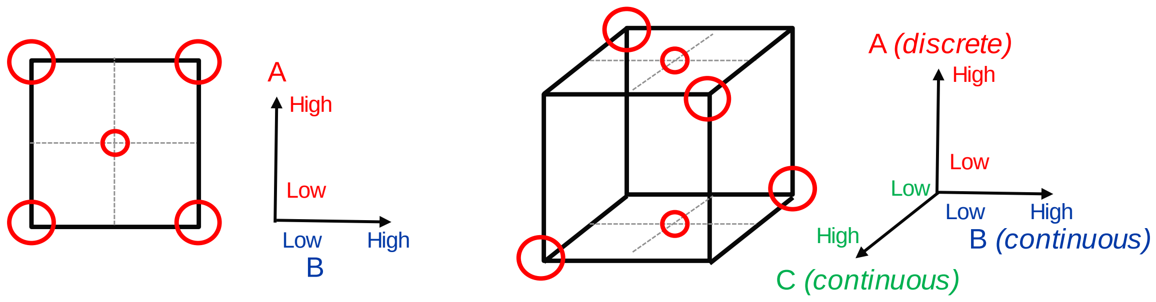


Figure 20: Factorial experiments with centre points (Left-to-right: 2^2 with centre point, 2^{3-1} with pseudo centre points)

The book 'Design and Analysis of Experiments' by Montgomery (Montgomery, 2009) provides a more detailed overview of the theory behind the design, implementation, and analysis of Factorial Designs and FFDs.

4.4.3 The Taguchi Method

The Taguchi Method (TM) for experimental design was introduced by Genichi Taguchi, who promoted a focus on quality as an "avoidance of loss to society" (Pignatiello & Ramberg, 1991). His Robust Parameter Design (RPD) techniques introduce a set of available orthogonal design arrays, which with high fractions, allow the user to investigate the effects of numerous variables at varying levels. The main strengths of the Taguchi method have been identified as being a practitioner friendly means of integrating statistical stools into the engineering process, and its emphasis on reducing undesirable variation in a product or process (Tanco et al, 2008). Furthermore, Pignatiello & Ramberg (Pignatiello & Ramberg 1991) presented a discussion of 10 aspects of the TM considered as 'triumphs', where its success in engaging a wider audience in quality improvement through experimental design.

However, much controversy exists around the design and analysis techniques promoted in the TM. Pignatiello & Ramberg also presented a review of 10 aspects of TM considered as 'tragedies', encapsulating much of the academic discussion of the techniques of the time (Pignatiello & Ramberg, 1991). More recently, Tanco et al reviewed the limitations of the TM, highlighting lasting concerns over both the catalogue of experimental designs, and the use of Signal to Noise ratio as the basis of analysis (Tanco et al, 2008). The highly fractional designs promoted in the TM result in main effects and interactions being confounded, meaning they should only be used when no interactions between variables is expected time (Pignatiello & Ramberg, 1991), while often being less efficient that fractional factorial designs with centre points in assessing both the effects of interactions, and linearity of effects (Montgomery, 2009).

Despite this controversy, the TM remains popular among practitioners, suggested by Ilzarbe et al as being a result of their practicality (Ilzarbe et al, 2008). Wu & Wu noted that the orthogonal design catalogue simplifies the designing of experimental programme significantly (Wu & Wu, 2012), while Rowlands, Antony & Knowles suggest that TM remains a 'useful starting point' for practitioners (Rowlands, Antony & Knowles, 2000).

4.4.4 Central Composite, Box Behnken, & Response Surface Methodology

While centre points can be added to FD & FFD designs to improve the understanding of linearity of effects, interpolation is still required to fully model the response of the system. Central Composite Designs (CCD) are based on FD or FFD designs, with the addition of both centre points and axial runs (Montgomery, 2009). An example of CCD based on a 2^{3-1} design is shown in Figure 21 with the additional points associated with the axial points shown in blue. The axial points are at set at a level based on a ratio, k, to the levels used for the FD/FFD treatment combinations, where a k=1 would correspond to the axial points being on the 'face' of the cube; such designs are known as Face Centred Designs (FCD). As such, CCD or FCD designs can provide additional insight into the system response across the entire experimental space.

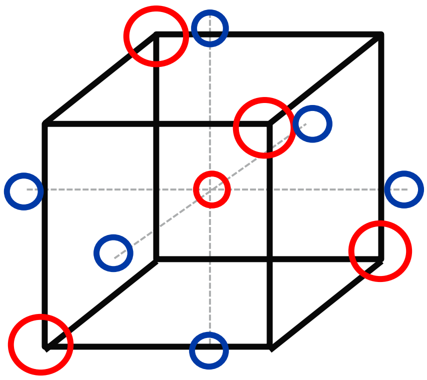


Figure 21: Central Composite design based on a 2^{3-1} design

Another means of characterising system response across the entire experimental space is through Box-Benhnken Designs (BBD) proposed by Box & Behnken (Box & Behnken, 1960). BBD use 3 levels for each factor, combining factorial designs with incomplete box designs, representing an efficient use of treatment combinations (Montgomery, 1990). An example of a 3 factor BBD is shown in Figure 22. BBDs do not include points in the 'corners' of the experimental space, and as such are useful when such treatment combinations are impossible or impractical (ibid).

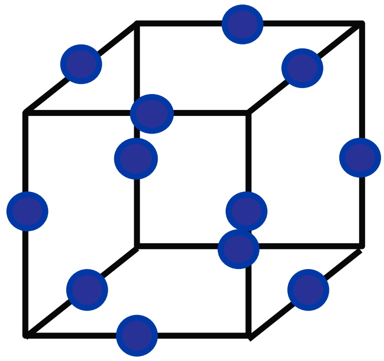


Figure 22: Box Behnken experimental design with 3 variables

Both the CCD and BBD designs are used as part of the Response Surface Methodology (RSM), which comprises a set of mathematical tools for analysing and optimizing the response of a system. Through characterising the entire experiment space, such designs allow for multidimensional equations to be derived that describe the response of the system to variations in each factor. RSM typically uses graphical techniques to visualise the response of the system, with the aim of predicting the optimal level of each variable (Montgomery, 2009).

4.4.5 DoE Limitations and Considerations for Use

The book 'Design and Analysis of Experiments' by Montgomery (Montgomery, 2009) provides a comprehensive overview of the considerations associated with the design, execution and analysis of experiments, providing a foundation for practitioners. Amongst the details, Montgomery emphasises the iterative nature of experimentation, suggesting that it is often to simplify experiments, while ensuring suitable levels of rigour are kept, encouraging the use of initial, low-resolution screening experiments to reduce the number of variables required for more detailed investigation (ibid). Montgomery also emphasises the role of expert judgement in the experimental planning stages in order to best select the variables to be investigated (ibid). When analysing experiments, Montgomery firstly recommends the uses of data visualisations, before then suggesting that Analysis of Variance (ANOVA) represents "probably the most useful technique in the field of statistical inference" (ibid). Finally, Montgomery reminds the experimenter to appreciate the difference between statistical significance and practical meaning, suggesting that while a factor or interaction may result in an effect that is significant compared to experimental error, it may not result in a response that is of any practical value (ibid).

In a review of both the literatures, and their own practical experience with DoE, Firka (Firka, 2011) presented a comprehensive list of issues that can influence the outcome of applications of DoE, as a result of deficiencies in planning, execution, and analysis. The list was grouped into three dimensions: the statistical dimension, the technological dimension, and the sociological one.

After their colleagues had identified only a small number of surveyed companies as using designed experiments, Tanco et al (Tanco et al, 2008) identified a total of 16 possible barriers to implementation of DoE through a literature review. A further survey with 101 respondents was then used to identify the main barriers as the 'low commitment of managers', and the 'poor background of engineers in statistics' (ibid).

Furthermore, it has been suggested that DoE methodology can be complex initially, and a poor comprehension of the methodology or the products/process can lead to deficient results (Booker, Raines, & Swift, 2001). To mitigate this risk, it is suggested that teams should be familiar with the product/process, and aware of the potential for misinterpretation of results (ibid).

Genichi Taguchi was among the most vocal critics of the classical DoE Methodology embodied by Factorial and Fractional Factorial designs (Tanco et al, 2008). While the TM has come under scrutiny for specific methodological elements, it is praised for being accessible to practitioners, and in turn, has promoted the inclusion of some of its concepts, such as variance reduction and clarification of process steps, into the classical methodology (ibid), while software capabilities have also lowered the previous barrier of knowledge of complex statistics (ibid).

4.4.6 DoE Applications

Prvan & Steel (2002) presented an annotated bibliography of cross-disciplinary papers published in a 5-year period in which fractional factorial experiments, and their derivatives, were employed. Of the 202 applications presented across 143 papers, only 25 came from the wider mechanical engineering field, constituting aviation, engineering, manufacturing, and materials (ibid), less than the biotechnology discipline alone.

Conducting a survey of 285 manufacturing companies in Spain, The Basque Country, and Baden-Wurtemburg, Tanco et al (2008) found that only 33% of respondents reported to know 'something' of DoE, while 23% had applied it to their problems in varying degrees of frequency. The survey highlighted that 71% of the failures reported in the application of DoE resulted from a lack of fundamental knowledge of the DoE methodology. In the same survey, 75% of the companies that responded use OFAT experimental programmes.

In a study of the literature published between 2001-2005, Ilzarbe et al (2008) reviewed the applications of DoE in the engineering field. In this review, 71% of the applications involved 5 or less design factors, and the most frequently used designs were those associated with the Taguchi method (31%), with Factorial and Fractional Factorial Designs representing a combined 30% of applications.

A literature review carried out as part of this research, concerning publications from 2000 onwards by the Society of Automotive Engineers (SAE) as a popular outlet for practitioners in the automotive industry, included a total of 167 papers with the key words 'Factorial Design', 'Taguchi Method', or 'Design of Experiments' as a cross section of practitioner literature. Of that cross section, ~90% of the applications used DoE in optimisation of a design or process, with prevalent sub-sets of design optimisation in the form of virtual testing using CAE tools, and in combustion optimisation. Both subsets allow for more resource efficient usages of high numbers of treatment combinations, and the use of space filling designs such as Latin Hypercubes is common. Only 5 papers identified within this cross section explicitly used DoE in the characterisation of a failure mode. In this cross section, the TM is prevalent, along with FD, FFD, and CCD designs.

4.5 Reliability & Accelerated Life Testing

Reliability can be defined variously as:

> *"The probability that the component or product will perform its intended function for a specific mission duration, under specified conditions" (Sarakakis et al, 2011)*
>
> *"The probability that a product does not fail under given functional and environmental conditions during a defined period of time" (Bertsche, 2008)*
>
> *"The probability that a unit will perform its intended function until a given point in time under specified use conditions" or the corollary "the probability that a unit will perform its intended function until a specific point in time under encountered use conditions." (Meeker et al, 2003)*
>
> *"Reliability is quality over time" (Condra, 1993)*

Regardless of the specific definition chosen, there are two clear points of note. Firstly, that reliability is time bound; when specifying or assessing reliability, it therefore must be done with reference to a usage duration, or 'mission' length. Secondly, that the usage conditions, with respect to both function and environment, are of critical relevance to the resultant reliability of the product.

While reliability is a significant criterion in the differentiation between automotive products, todays OEMs are under increasing market demand to bring higher technology products to market, with a reduced time and cost to market (Escobar & Meeker, 2006) (Yadav et al, 2006). While OEMs previously focused on improving manufactured quality alone through inspection and quality control, increasing reliability is achieved through improvements to design and manufacturing capabilities (Escobar & Meeker, 2006). This has motivated changes in the product design and development process, such as concurrent engineering and experimental design, with more upfront testing of materials, components, and systems (ibid). The increasing reliability of products has in turn made reliability testing a more challenging task (Yadav et al, 2006), that represents a significant proportion of the total product development cost (Elsayed, 2012).

While reliability can be influenced through the life cycle of a product, the choices made in the initial design stage have a significant effect on the final reliability achieved and the resource requirements required to do so. Appropriate understanding of the usage conditions and the definition of appropriate reliability requirements will dictate many of the design and development activities of a firm engaged in NPD (Sarakakis et al, 2011). However, it has been suggested that there is often a gap between what engineers think the customer does, what the customer themselves think they do, and what the customers actually do (Campean et al, 2002). A comprehensive integration of reliability engineering into NPD is required to provide a systematic procedure to the NPD process and highlight potential critical features in the product design (Yadav et al, 2006). Appropriate product testing early in the product design stages is considered crucial in achieving the final reliability requirements (Elsayed, 2012).

Reliability is most accurately assessed through testing components in conditions equivalent to end usage. For products with long life requirements, this requires long test durations, and to demonstrate statistical confidence in results, an extensive number of samples (Elsayed, 2012). This problem is compounded when products are designed to have very high reliability where it is expected that few units will fail or degrade in use level testing of practical durations (Escobar & Meeker, 2006). As it is therefore costly and impractical to directly assess reliability in such a manner, reliability engineers utilise methods to 'predict' the reliability of a product at use level using knowledge gained from testing at conditions other than normal operating conditions such as to induce failure or degradation in a shorter test duration, known as Accelerated Life Tests (ALT) and Accelerated Degradation Tests (ADT) respectively (Elsayed, 2012). In order to identify the critical failure modes for predictive testing, engineering judgement knowledge, in combination with warranty performance of similar products and assessment methods such as Failure Mode and Effect Analysis (FMEA) are employed (Yadav et al, 2006).

A paper by Sarakakis et al (Sarakakis et al, 2011) provides a comprehensive summary of the considerations associated with a Design for Reliability (DFR) process. In their review, Sarakakis et al suggest that many DFR challenges in industry result from a lack of appropriate knowledge of the functional and environmental conditions associated with usage, suggesting upfront work is required to define the full range of suitable conditions to be considered in testing. They also suggest that the product life duration and conditions should be captured in a set of well-defined and meaningful requirements, agreed with all stakeholders in the NPD process. As those requirements structure many of the product design and development activities, and therefore resource requirements, their validity is of utmost importance.

Sarakakis et al (Sarakakis et al, 2011) go on to suggest that appropriate and thorough effort should be applied in the design of ALT plans, as deficiencies in the design stage can lead to failure in creation of required product knowledge. An understanding of failure mode behaviour is therefore required in order to select the appropriate ALT models, and DOE is considered best practice in the screening of the most significant stresses.

4.5.1 Toolset for DFR

The book “Designing Capable and Reliable Products” (Booker et al, 2001) provides an overview of the key tools and techniques associated with DFR in NPD, including FMEA, QFD, and DFM/DFA, describing such tools as generally being team-based design evaluation tools that should be used from concept design through to detailed design of the final product.

Failure Mode and Effects Analysis

Failure Mode and Effects Analysis (FMEA) is a structured tool from the Design for Reliability methodology that allow product engineering teams to identify potential failure modes, assess the risks associated with their instance, and rank them such that the team can subsequently identify and implement corrective actions in a prioritised manner. Booker described how FMEA is a key input into the design of capable and reliable products which should be started as soon as possible in the NPD process, and that effective application can reduce late design changes and costs associated with product failures (Booker, Raines, & Swift, 2001). Henshall, Campean, and Futter (2014) described how FMEA is generally considered as a key tool in industrial applications of NPD.

However, Henshall, Campean, and Futter (2014) described how FMEA is not always employed most effectively in complex sociotechnical contexts, identifying the key challenges associated with the application of FMEA as: system complexity, systems deployment, multidisciplinarity, and multi-domain integration. Booker et al (2001) also identify further issues, including its often subjective nature, timing within a NPD programme, and the importance of meaningful input from the customer and suppliers.

Quality Function Deployment

Quality Function Deployment (QFD) is a tool from the Design for Six Sigma methodology, taking the form of a planning matrix through which the customer's requirements can be transformed into product design characteristics, process requirements, and control standards (Bralia et al, 2007). While FMEA champions the role of the engineer, QFD champions the customers' needs (ibid). Booker et al (2001) identifies the benefits of QFD to be associated with improved customer satisfaction and product quality, while potential issues include the multidisciplinarity of the required team, its often subjective and procedural nature, and a tendency to not fully complete.

Team based problem solving methods

A number of team-based problem-solving methods exist, emanating from a number of different methodologies, including 8 Disciplines (8D), Plan-Do-Check-Act (PDCA), and Design-Measure-Analyse-Improve-Control (DMAIC). In general terms, the methods each emphasize a robust definition of the problem, exploration of the underlying causality though RCA, and ensuring that the corrective actions implemented permanently resolve the problem through ongoing control.

The 8 Disciplines (8D) method is outlined by Sarakakis et al (2013), providing an overview of the 8 steps (or 'disciplines') of the process, and describes how it has become a standard within the Automotive Industry.

The DMAIC method is a key component of DFSS methodology, providing a structured approach to understanding and addressing an identified problem, and is outlined by Sokovic, Pavletic, & Pipan (2010) who summarise the key to its implementation as being able to measure, and express the product or process in data.

The PDCA cycle, also known as the Deming Wheel, is a process for product or process improvement, and is also outlined by Sokovic, Pavletic, & Pipan (2010). The PDCA method emphasises action learning through iteration, repeated until an effective solution has been implemented (Rothwell & Sullivan, 2007).

Root Cause Analysis RCA

Root Cause Analysis (RCA) is a collective term for several available tools to explore the causality of a problem. Such tools include Fault Tree Analysis (FTA), DoE, Ishikawa Diagrams, and 5 Why, and the primary objective of each is to identify the root cause(s) of a problem, that if corrected will resolve the failure of a product or process.

FTA is an established RCA tool that utilises a graphical representation of a system, from a deductive perspective, that models the possible root cause(s) of a failure mode in a top-down manner and is defined in EN61025. The observed failure mode would be defined as the top of the 'tree' with the separated downward branches being the possible causes, with each branch continuing to propagate further sub-branches until independent potential causes are identified. Potential root causes are then eliminated through investigation, until only the actual root cause(s) remains.

Ishikawa Diagrams, also known as fishbone diagrams, are another RCA tool that utilises a graphical representation of causality. With the failure mode as the 'head' of the fish, potential causes are grouped into the 'bones' by category dependant on the context. For product failures, those categories may be subsystems, while for processes, the categories may be major sources of variation. Each bone is then further explored through sub-branches in a similar manner to FTA.

5Why is an RCA tool typically employed in both DFR and DFSS methodologies, representing one of the simplest RCA tools, requiring little statistical or graphical analysis. The method calls for the problem-solving team to repeatedly ask the question "Why?" when presented with a problem, stripping away the symptoms until a true root cause is identified. The rule of thumb associated with the tool is that five iterations of the question "Why?" will be enough to identify a root cause, although it may require fewer, or more than five iterations depending on the complexity of the problem.

Shainin Methodology

Steiner et al (2007) provides a comprehensive overview of the Shainin System™ as a quality tool describing the core principles of the methodology as the use of the Pareto principle to reduce the number of possible variables, an emphasis on additional product level experimentation or observations, and the search for the RedX, which would be either a single variable, or an interaction of variables. This overview goes on to provide a description of several Shainin System™ tools, including paired comparison, describing how experiments or observations associated with the extremes of the populations as BOBs (Best-of-Best) and WOWs (Worst-of-Worst) are used to reduce the number of possible RedX variables.

Tanco et al (2008) provides a summary of additional literature on the Shainin System™, focusing on its approach towards experimentation, while describing that the intellectual property stance of its creator has led to it relatively infrequently being discussed in literature.

4.5.2 Accelerated Life Testing Methodology

Accelerated Life Testing (ALT) methodology is a key element of DFR. The fundamental objective of ALT is to quantify the life characteristics of a product at use level in the form of a Probability Distribution Function (PDF), with reduced time and resource requirements (Vassiliou & Mettas, 2003). This is achieved through accelerating failures in a controlled, repeatable, laboratory environment through the application of accelerated stresses (ibid). The knowledge gain from ALT can then be used to predict warranty returns, perform product or process risk assessments, or to compare alternative designs, and is ultimately validated through the observation of field results (ibid).

Acceleration has numerous definitions within DFR, but generally relates to obtaining reliability data quicker (Escobar & Meeker, 2006). Product failures in controlled tests can be accelerated in several different manners, including usage rates, chemical ageing, stress amplitudes, and a combination thereof (ibid).

ALT tests can be Quantitative, Qualitative, or Environmental in nature (Vassiliou & Mettas, 2003). In order to design a suitable accelerated test to determine a quantitative model of reliability, product failure modes and the conditions that accelerate them should be known (Escobar & Meeker, 2006) or characterised through DOE (Vassiliou & Mettas, 2003). When using severe stresses associated with qualitative testing, any sample failures should be confirmed as representative for the results to be valid (Escobar & Meeker, 2006).

Reliability Demonstration Tests (RDT) completed on pre-production samples, demonstrate that the product meets the reliability and confidence requirements for production, while Reliability Growth Tests (RGT) provide continuous feedback on the reliability improvements to a product during the product design and development phases of the NPD cycle (Elsayed, 2012).

The ALT design process has 3 main steps: definition of test objectives; determination of accelerated stress(es) and the model(s) that will be used to analyse results; and determination of stress profiles and sample allocations (Elsayed, 2012). The acceleration models selected to describe the underlying stress-life relationship for a given ALT, can either be based on known failure physics (or chemistry), or on empirical knowledge (Escobar & Meeker, 2006). Empirical models require significant justification and evidence to be supplied when used to extrapolate reliability predictions at use level, and improper assumptions may require in incorrect conclusions (ibid).

In order to generate representative failure data, and therefore valid predictions of product reliability, accelerated test must generate the same failure modes witness in field running, and accelerated variables should be chosen to correspond with variables associated with usage (Escobar & Meeker, 2006). In the event of uncertainty around the effects of variables, sensitivity analysis, through probe tests or Design of Experiments methodology should be employed (ibid).

ALT programmes should be designed to minimise the requirement for extrapolation, and increase the confidence associated with any predictions of reliability data at use level (Escobar & Meeker, 2006). Reliance on testing at one accelerated stress level only can be insufficient unless the stress-life relationship is well understood, and appropriate justification can be provided (Vassiliou & Mettas, 2003).

In addition to design considerations associated with the accuracy and validity of an ALT programme, it is typical to also be presented with restraints on resource (Elsayed, 2012). Experimental test durations will influence the cost of an ALT programme and may impact on the physical capacity of a business, while the number of test samples required, typically at higher unit costs than the resulting product, will also influence the cost of a programme (ibid). Therefore, it is possible that a business engaged in NPD may be required to balance the resources required with an ALT programme against the potential warrantee exposure associated with the accuracy and validity with which the reliability of the product can be predicted.

In the case of high reliability products, Accelerated Degradation Testing (ADT) can yield more valuable life data than test to failure ALT tests, in the form of knowledge of wear progression and failure physics. Assessing high reliability products can prove difficult as the high reliability associated with the product makes it highly resource intensive to generate a sufficient number of failures in a reasonably short time, and the high unit cost associated with such products can make it unfeasible to collect enough data for classical analysis (Erto & Giorgio, 2002). ADT necessitates a method for assessing the performance of a product at different time intervals through non-destructive observations or inferences. As such, ADT requires highly controlled empirical test facilities (Doikin et al, 2018).

4.5.3 ALT limitations

In the application of accelerated stresses, it is possible to introduce unrepresentative failure modes, and the extrapolation associated with the analysis process places significance on the selection of appropriate numerical models (Escobar & Meeker, 2006). As such, in order to best ensure the successful implementation of an ALT programme, the design and execution of the test programme should be completed by a cross functional team (Escobar & Meeker, 2006). It is best practice for that team to include; individuals with knowledge about the product and its use environment, experts with knowledge of the failure physics, and statistical experts with knowledge of the design and analysis of reliability testing (ibid).

A comprehensive summary of the potential limitations associated with ALT is discussed by Meeker & Escobar (Meeker & Escobar, 1993) and later revisited by Meeker et al (Meeker et al, 2013), identifying considerations associated with the design, execution, and analysis of ALT programmes. A recurring theme through those limitations concerns the existence and availability of knowledge pertaining to the product usage and its potential failure mechanisms, where deficiencies in such knowledge can limit the effectiveness of the ALT methodology. Furthermore, it is suggested that a common limitation is application is the underutilisation of degradation data, which for the same test duration, will always provide more knowledge of the product and its failure modes than just time to failure data, particularly for high reliability systems (Meeker et al, 2011).

4.5.4 Assessing and correlating field reliability

Design for Reliability does not stop when a product enters production (Sarakakis et al, 2013). Knowledge gained from field reliability verifies the results of any ALT programme, allowing correlation and subsequent refinement of models and predictions, whilst also providing a suitable metric for measuring the success and efficiency of a PV process. While tracking and managing of warranty returns is in itself a routine and often stressful task, the reliability knowledge associated with field reliability is considered an important subset of organisational knowledge that needs to be managed effectively for the long-term benefit of the business (ibid). It is therefore best practice to employ an integrated system through which field reliability is managed in the same method, utilising the same knowledge banks, as is used in NPD (ibid).

Failure reporting, analysis, and corrective action systems (FRACAS) are typically software-based tools used in industry for the management of information relating to product failures in a standardise, and indexable systems. Sarakakis et al (Sarakakis et al, 2013) describes the considerations associated with designing and implementing a FRACAS system, describing how such a system can be used to improve product reliability through design iterations, and from field experience.

4.6 Summary

This Chapter began by discussing the role of Knowledge, and Knowledge Management in NPD, describing it as a differentiator between organisations, influencing business performance. The role of Expert Judgement in the face of uncertainty in the NPD process was introduced, alongside how Expert Elicitation can be used to access that tacit knowledge. The Delphi Method was then proposed as a flexible technique for Expert Elicitation, with best practices for effective applications identified.

Systems Modelling was then introduced as a methodology that provides frameworks that can improve the understanding of complex systems, acting as a means for transforming uncodified knowledge into useful boundary objects in the NPD process. Causal Loop Diagrams were then presented as a form of System Model that represent a simple language for communication between the different audiences in an NPD programme, while being resource sensitive to generate and develop.

The role of experimentation in NPD was then introduced, while it was presented that in industry, experimentation was often performed in a sub-optimal manner. Experimental Design Methodology was then discussed as a key part of the scientific method, identifying how it can present the most efficient use of experimental resources. The 3 basic principles of experimental design have been identified as randomisation, replication, and blocking. A selection of the available experimental design methods has been introduced, including Factorial and Fractional Factorial design, and the limitations and considerations associated with their usage have been discussed as being 3 dimensional: statistical; methodological; and sociological. Finally, a survey of applications of experimental design has identified only 23% of respondents utilising the methodology in European industry, while a review of automotive practitioner literature showed that the significant majority of applications were in design or process optimisation, and only ~4% of the applications explicitly used DoE in the characterisation of failure modes. It was also presented that in industry, experimentation is often performed in a sub-optimal manner.

Finally, the concept of Reliability and Accelerated Life Testing was presented, demonstrating the typical methods used in NPD programmes to validate that a product meets all of the customer’s requirements, in a controlled, and resource efficient manner. The challenges and limitations associated with the methodology were discussed, demonstrating that while ALT can generated significant and timely knowledge about a product, it is in part reliant on Expert Judgement to mitigate uncertainties about a products performance, usage, possible failure modes, or otherwise.

The next Chapter provides focal theory in each domain by exploring the existing state of each domain within Delphi Technologies at the time of writing this Thesis.

Chapter 5 Existing Technology, Processes, Tools, & Methods Within Delphi Technologies

5.1 Introduction

This Chapter presents an overview of the sociotechnical context of this research. A new, and comprehensive overview of the pre-existing state of Knowledge Management and the use of Expert Judgement and Expert Elicitation in the Delphi Technologies NPD process will be presented, identifying parallels and gaps to the practices identified in Chapter 4. Furthermore, the methods employed by Delphi Technologies for validation of the Euro VI family of products, including both accelerated test design and correlation with field results across applications, will be presented. This new analysis will discuss the limitations associated with the process, through the propagation of uncertainty associated with knowledge loss across the existing boundaries associated with field failures.

5.2 About the author

At the time of writing, the author of this paper was an employee of Delphi Technologies within the Product Validation department, representing an industrially based research engineer with links to both academic institutions associated with the research. When this research commenced, the author had already been an employee of Delphi Technologies for 4 years, after joining with experience elsewhere in the automotive industry. Prior to this research, the author supported validation of the first of the F2 family of products to reach production, before moving to support change validation programmes.

Within Delphi Technologies, the Product Validation department is responsible for demonstrating that new products meet their associated requirements, including reliability, to both the business and its customers. The validation department is therefore involved in all stages of the NPD process, from evaluation of a new opportunity, to supporting field issues in high volume production, and in doing so, interfaces with every element of the business alike.

While completing this thesis, the author was a Senior Engineer in the Diesel FIS Validation team, responsible for Test Development across all market sectors, supporting product development programmes for passenger car, and commercial vehicle platforms alike. In this role, the author built on his previous experience in the business, alongside the knowledge associated with this research, to design empirical test programmes to support both new and existing products.

5.3 Injector performance characterisation at Delphi Technologies

The performance characteristics of DFI and observed using hydraulic test rigs, with Fuel Metering Units (FMU) as the key element of a suite of measurement systems. This section will provide an outline of such facilities utilised by Delphi Technologies, provide detail of the process used to establish control of this measurement system, and the data made available through the performance characterisation process.

5.3.1 Injector performance characterisation testing

The injector samples performance will be characterised on a single cylinder hydraulic rig, hereby referred to as the 'performance rig'. The performance rig utilises a Delphi Technologies distributed high pressure pump as a pressure source to provide pressurised fuel to a common rail. That common rail then serves as a pressure source for a single injector, installed in a bespoke single injector pocket. The pocket dimensions are the same as those on the cylinder head of an engine, and an injector clamp is tightened such as to provide the specified axial load for application.

To characterise the performance of the injector across an operating range of rail pressures and injection logics, a Fuel Metering Unit (FMU) is integrated into the pocket of the performance rig. A comprehensive review of the techniques used to meter injected fuel presents the different options available in FIE development and production (Lillington, 2015). The performance rig used in this experimental study uses an Akribis FMU supplied by Loccioni which infers volume of injected fuel by measuring pressure and temperature of the measuring chamber, and the travel of a nitrogen backed piston using an inductive displacement transducer. In doing so, the Akribis FMU provides the quantity of injected fuel per event and a high-speed injection rate measurement. Through the capture of repeated injection events at a given condition, in addition to individual values for each repeat, the Akribis FMU can also determine the repeatability of an injector in the form of average and standard deviations in quantity. The Akribis FMU used in this experimental study has a measuring range of 0.8mm^3 to 350mm^3 with a ±0.25mm^3/stroke accuracy between 0.8-250mm^3/stroke, and ±0.1% for 250-350mm^3/stroke. The Akribis FMU provides a compact, direct measurement of injected fuel volume, with suitable accuracy across the full operating range of typical Euro VI applications.

The injector performance characterisation test is largely automated, with a highly controlled operating process to ensure consistency and validity of results. However, some elements remain under the potential influence of the operator, so a rigorous quality procedure is associated with the operation of the performance rig. That procedure includes ensuring consistent fluid, and injector pocket, temperatures are reached before injection measurement beings, and that negligible leakage from the measurement chamber is measured.

Once the operator has followed the process to ensure that the rig is thermally stable, with adequately controlled leakage from the injected fuel volume, an automated test plan is initiated. For the injectors used in this experimental study, that test plan involves characterising the performance of the injector at 5 different system operating pressures, ranging from the natural pressure of the system associated with idling, and the rated pressure of each application. The test plan first raises the system operating pressure to the rated condition, before allowing a period of dwell sufficient to best ensure thermal stability, and highly stable rail pressure control. The test plan then captures a specified number of observations at each point in a reducing sweep of injection logics, characterising the full range of typical injection quantities. The logic sweeps vary in separation, with the ballistic region of the gain curve characterised in reduced logic separations than used for the linear portion of the gain curve. Once the logic sweep has been completed for a given system operating pressure, the test plan reduces the pressure to the next lower pressure, and again allows for thermal stability and rail pressure stability to be fully achieved. By performing multiple observations at each combination of rail pressure and logic, the performance rig characterises the performance of an injector with respect to both the average fuel quantity associated with each point, and the corresponding shot-to-shot variation.

5.3.2 The fuel gain curve

The performance of an injector is primarily characterised through its fuelling gain curve. The fuel gain curve provides a visualisation of the average fuel quantity injected for a given logic demand and is typically repeated for the different system operating pressures associated with the injector test plan. The fuel gain curve represents the average of 5 injections at each combination of logic and system operating pressure. An annotated gain curve shown in Figure 23.

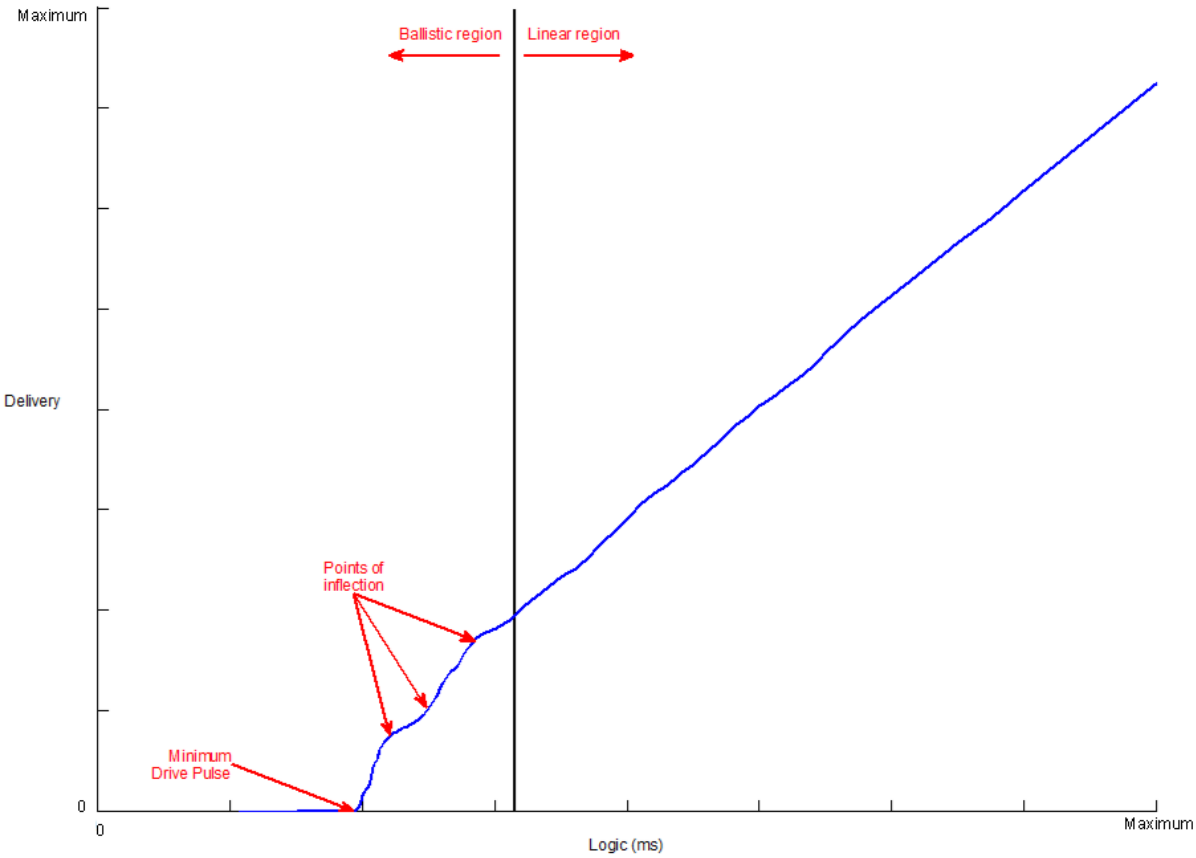


Figure 23: Annotated DFI fuel gain curve

For common rail fuel injectors, the gain curve is typically divisible into two distinct regions. The first such region is known as the ballistic region, where the injection logic is such that the end of injection occurs while the NCV is not in a state of equilibrium with respect to its mechanical position, the hydraulic condition of the control volume, and the electromagnetism associated with the stator, as described in §3.2.2. The lower bound of the ballistic region is the Minimum Drive Pulse (MDP) and is characterised by the end of injection taking place in varying states of mechanical, hydraulic, and electromagnetic conditions.

In the lowest portion of the ballistic region close to MDP, the nozzle needle is still in motion at the end of the injection logic, and the delay in it returning to its bottom seat is dependent on the motion of the nozzle at that instant. At higher logics within the ballistic region, the nozzle needle will have reached its top seat, but the pressure in the hydraulic control chamber would yet to have stabilised, resulting in variations in control chamber depressurisation rates at the end of injection, and thus influencing the total quantity of injected fuel. Injection quantities in this region of the gain curve for different injectors are therefore dominated by the performance of the NCV.

The ballistic region of the gain curve is characterised by a fuel gain curve with varying gradients in the injected fuel quantity for increasing logic. This results in points of inflection in the gain curve, where an increase in logic does not result in a proportional increase in injected fuel quantity. This region is associated with the significantly reduced injection quantities associated with pilot and post injections, and part load running.

Once injection logics are of enough duration that the end of injection occurs with the NCV in a state of equilibrium, the fuel gain curve linearizes for increasing logics. In this region, both the NCV and nozzle have reached their respective top seats, and pressure stability is achieved in both the control chamber and the nozzle barrel volume alike. This region of the fuel gain curve is known as the linear region, and differences from injector to injector are no longer dominated by the performance of the NCV, but instead by the performance of the nozzle itself.

In addition to characterising the performance of an individual injector, the fuel gain curve can be used to compare the relative performance of multiple injectors, or to compare the performance of an individual injector over time. For a given design level of injector, or for a population of injector samples, it is typical to identify a nominal fuel gain curve, hereby referred to as the 'nominal', alongside limits of acceptable deviation of samples away from that nominal curve. The nominal is utilised by the OEM to determine fuelling maps for the engine, with a level of part to part variability acceptable as defined in the product specification, and that part to part variability is then used by Delphi Technologies to determine the pass-fail criteria for injector deliveries. It is typical for the deviations from the nominal to be defined in fuelling quantities for given logic intervals, with reduced fuelling tolerances at smaller injection quantities. For a given nominal, it is possible to assess the performance of injector(s) against that tolerance banding through observing the difference, in terms of injected fuel quantity, between the injector(s) and the nominal.

In order to assess this difference in performance from the nominal, the fuel gain curves are typically 'trimmed' such that the minimum drive pulse for each aligns. The MDP for a given injector will vary within a small range dependant on several factors associated with design and manufacture. To apply an MDP trim to a number of injectors, the gain curves are typically aligned at a logic that corresponds to 1mm^3, such that it is within the capability range of the Akribis FMU.

The fuel gain curve can also be used to assess the performance of a single injector over time. In order to do so, the difference in terms of injected fuel quantities for subsequent gain curves is determined through interrogation of the performance test results. The typical visual representation of fuelling change over time, or Drift over Time (DoT), consists of an overlay of the fuel gain curves for each measurement interval, along with a differential line plotted against a second y-axis that represents the fuelling change between the fuel gain curves. An example of this visualisation is shown in Figure 24.

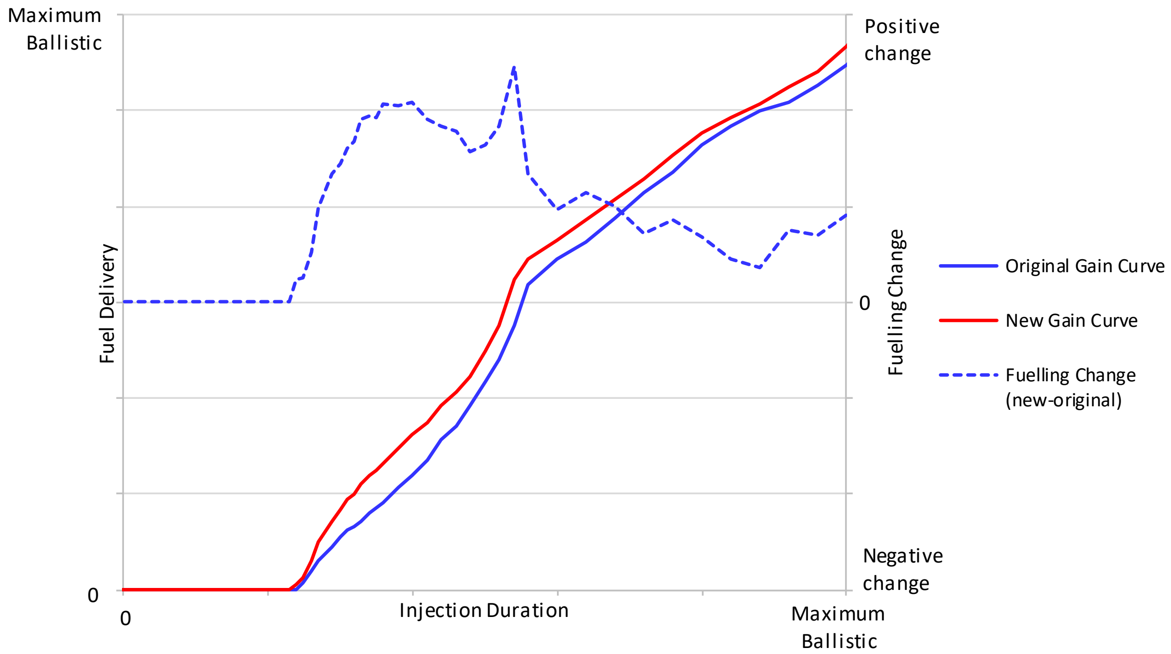


Figure 24: Visualisation of fuelling change between tests for a single injector

In addition to providing a visualisation of the performance of an injector over time, the results of a injector performance characterisation test can also be used to provide a quantification of the magnitude and direction of that change. Typically, injector DoT is determined through the interrogation of the fuelling gain curve at the rated pressure of the FIE system only. While it is possible to quantify injector DoT at any system operating pressure, its effects are largely proportional, and most severe at the highest system operating pressure of the FIE system. An example of this is shown in Figure 25 where the drift curves for 5 different injection pressures are shown for the ballistic region of the gain curve.

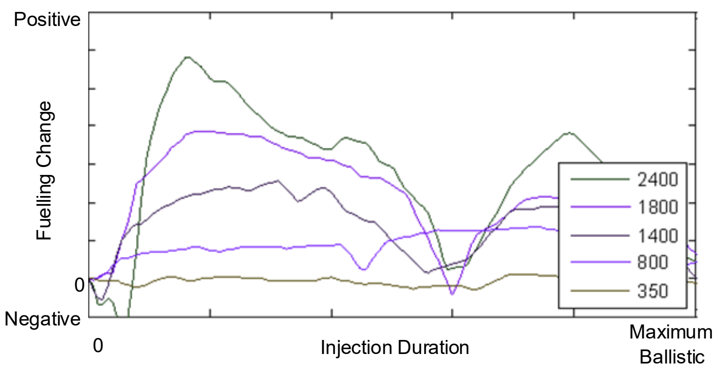


Figure 25: Fuelling drift for a single sample at given values of rail pressure (in bar)

The specific form of that quantification has varied between the customer projects in alignment with the customers' requirements. For some projects, maximal values of fuelling change, expressed in quantities or percentages, are reported against specific points, typically concerning injection quantities or durations associated with multiple injection strategies. For other projects, that quantification has taken the form of a maximal value of fuelling change within a range of fuelling quantities, known as fuelling zones, an example of which is shown in Figure 26 where two zones describe the ballistic region of the gain curve while a third concerns long logic portion of the curve. Alternatively, for some projects, the quantification of injector DoT concerns the maximal value of fuelling change across the full fuel gain curve.

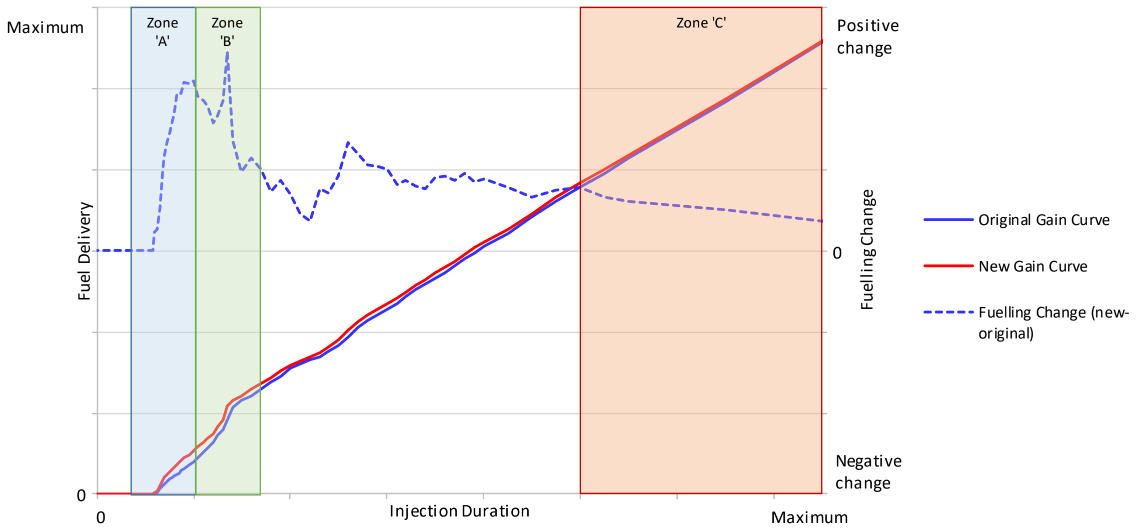


Figure 26: Example of fuelling zones used to quantify fuelling drift

The analysis for fuel DoT can be performed on gain curves without MDP timing correction (MDP Trim), or on gain curves trimmed such their MDPs align. The un-trimmed gain curve represents the DoT associated with both changes in the timing of the injector, and any wear modes that influence the shape of the gain curve. Comparing trimmed gain curves allows the analysis of DoT to be isolated to only those wear modes that influence the shape of the gain curve. In the ballistic region of the gain curve, such wear modes typically concern the NCV, while DoT in the linear region of the gain curve is typically associated with wear modes concerning the nozzle.

5.3.3 Shot-to-shot variability

Whilst interrogation of the fuelling gain curve(s) represents the primary means for characterising the performance of an injector, several secondary methods are also typically employed. The first of these represents the repeatability of the injector, typically expressed through the quantification of the ±3SD fuelling range associated with the injected fuel quantity. As the injector performance is characterised through the sampling of 5 or more injections at each system operating pressure and injection logic combination, it is possible to express the shot-to-shot variation curve associated with an injector in addition to its average fuel gain curve. Due to the variations in mechanical, hydraulic, and electromagnetic conditions associated with the ballistic region of the gain curve, Euro VI injectors typically exhibit more sensitivity to shot-to-shot variation at smaller fuelling quantities.

5.3.4 Injection timing

Another secondary method for characterising the performance of an injector is through an interrogation of the metrics that describe its timing. There are three timing metrics typically used to characterise the performance of a Euro VI injector, all express with the units of microseconds: Injector Opening Delay (OD), Injector Closing Delay (CD), and NCV Valve Closing Delay (VCD). Figure 27 visualises the SOR and EOR in terms of the injector drive waveform and the rate of fuel injection.

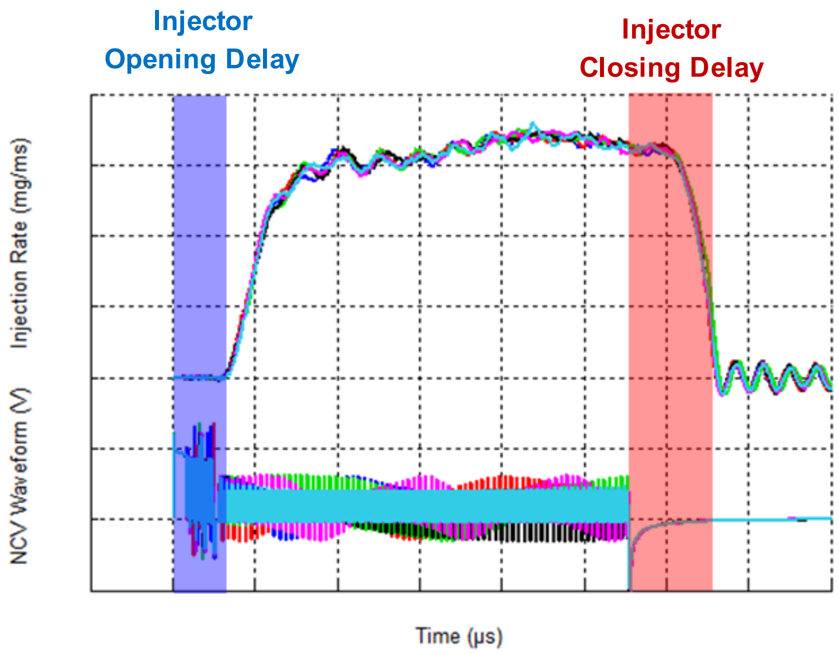


Figure 27: Visualisation of opening and closing delay of a DFI

The SOR timing metric describes the time delay between the start of the injector drive logic and the start of measured injection rate captured on the injector performance rig. Similarly, the EOR timing metric describes the time delay between the end of the injector drive logic and the end of measured injection rate captured on the injector performance rig. Both the SOR and EOR metrics are typically expressed as curves complementary to the fuelling gain curve, visualising how the timing of the injection vary with different injection logics for a given system operating pressure. For typical Delphi Technologies Euro VI injectors, the SOR for an injector is typically constant for a given operating pressure, but comparison of the SOR timing for multiple injectors, or for a single injector over time, can inform the understanding of any performance variation. For typical Delphi Technologies Euro VI injectors, the EOR for an injector varies throughout the ballistic region of the gain curve before then becoming constant for long logic durations. As such, the EOR curve for an injector provides additional insight into an injector's fuelling gain curve, and additional comparison between multiple injectors, or for a single injector over time, can inform the understanding of any performance changes.

The VCD metric characterises the closing behaviour of a 3-way NCV, and in conjunction with Injector CD, the associated hydraulic delay between NCV closure and the end of injection. The VCD is derived through interrogation of the injection drive voltage as captured on the injector performance rig. For a given injector, the change in drive voltage over time, expressed as Δvolts in volts per microsecond, can be calculated. When the NCV valve assembly, including the armature, returns to its bottom seat, a back-EMF can be observed in the Δvolts measurement, providing a means for characterising the NCV closure behaviour on an injector for varying injection drive logics. The hydraulic delay, which describes the delay between the NCV reaching its bottom seat, and the nozzle closing, can then be inferred from the Injector CD and VCD metrics. An increase in NCV closing delay indicates an increase in the travel time of the NCV valve pin assembly, and/or a change in valve closing speed. The VCD curve for a typical Delphi Technologies Euro VI injector will show variation in the ballistic region of the gain curve, as the end of injection takes places in a variety of mechanical, electromagnetic, and hydraulic conditions.

5.3.5 Injector leakage

The final secondary method for characterising the performance of an injector is through the assessment of its leakage when exposed to high pressure fuel. Leakage of a typical common rail can be differentiated as being either static leakage, or dynamic leakage, and is most typically associated with NCV.

Static leakage describes any leakage through clearances in the NCV, not associated with actuation of the valve. In pressure balanced NCVs, the valve pin-to-guide clearance is exposed to pressure at all conditions and is therefore a source of static leakage. A simplified representation of static leakage in the 3-way valve concept is shown in Figure 28.

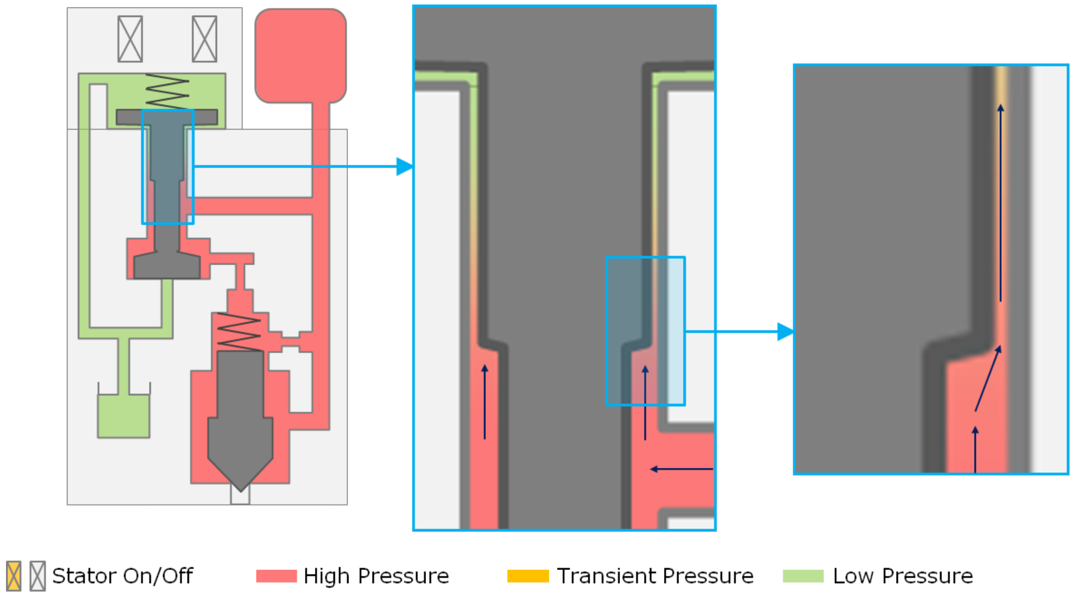


Figure 28: Static leakage in 3-way valve concept

Dynamic leakage describes fuel leakage associated with the actuation of the NCV. A simplification of dynamic leakage in a 3-way NCV concept is shown in Figure 29. In such a concept, high pressure volumes are ultimately isolated from the low pressure drain minimising dynamic leakage by the valve, while in other concepts, dynamic leakage is proportional to injection duration and pressure.

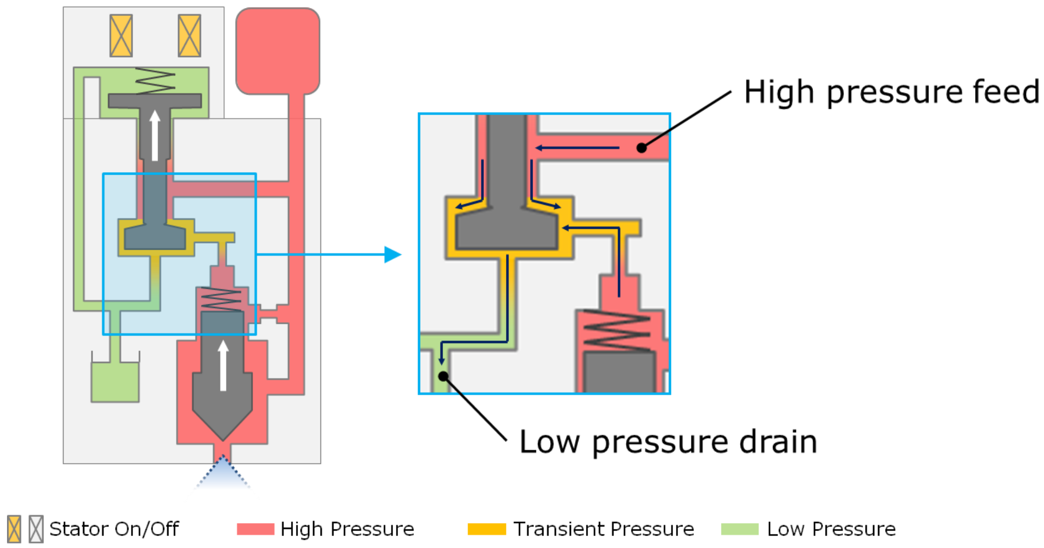


Figure 29: Dynamic Leakage in a 3-way valve concept

The static and dynamic leakage of an injector can be used as an acceptance criterion for as new injectors, or to compare the performance of multiple injectors or for a single injector over time. As the leakage of the injectors forms a significant element of the efficiency of a common rail FIE system and therefore for an engine and vehicle, it is typical for the OEM to specify limits for both individual injectors and the whole system in the 0km condition. When comparing the leakage of an individual injector over time, a change in static leakage is typically associated with either wear to the guided portion of the valve stem, resulting in an increase in the effective annular area exposed to high pressure, or to gross wear to the NCV bottom seat symptomatic of significant hard particle debris damage. A change in dynamic leakage over time for a single injector is typically associated with a change in either the effective lift of the NCV valve assembly, or to the motion of the NCV pin between seats. Wear to the NCV valve seats can result in an increase in the effective lift of the valve, increasing the period in which the low pressure drain orifice is exposed to high pressure fuel. Alternatively, wear can result in changes in the hydraulic conditions which govern the force balance acting on the valve assembly, and therefore motion of the valve, again potentially increasing the duration of time in which the drain orifice is exposed to high pressure

5.4 Accelerated Life Testing methodology within Delphi Technologies

5.4.1 Identification of potential failure modes associated with a product design

Delphi Technologies utilise several well-established DFR tools and methods within the existing PV process. After identifying a Product Design Specification (PDS) intended to meet the customer requirements, a cross functional team led by PV conducts a product application review. In this review, the PDS is reviewed alongside information from the customer pertaining to anticipated duty cycle(s), and the results of initial analytical experimentation using simulation tools. In parallel, the Design FMEA (DFMEA) is completed, identifying potential product failure modes. Using the results of the product application review, and the DFMEA, the PV function then completes a risk matrix, in which empirical tests are identified and prioritised for each failure mode, such that a validation plan can be generated. That validation plan would typically incorporate reliability demonstration on hydraulic test rigs, alongside more specifically targeted ALT tests, and testing using customer engines and/or vehicles.

The DFMEA is then revisited throughout the NPD process such as to demonstrate a reduction in product risk through identification and completion of empirical and analytical experiments through the course of a programme. Furthermore, DFMEAs for historic products or applications should be reviewed in new programmes as appropriate. However, as perhaps typical for many lean

Potential product failure modes are explored by cross functional teams in DFMEA, PFMEA, and IFMEA activities that are reviewed through the entire NPD process. However, as identified in §4.5.1, FMEAs are not always effectively within a complex sociotechnical context.

5.4.2 Design of empirical testing programmes

Accelerated Life Testing methodology is used extensively, both for prediction of system reliability at use level, and for assessing the robustness of products to specific failure modes through targeted testing. ALT programmes are designed by the Validation Engineers assigned to each project, with the support of cross functional teams, and are then approved both by the business and the customer. Some subsystems are supplied by third parties, who are in turn responsible for the validation of those subsystems, as approved by both Delphi Technologies and its customers. Delphi Technologies retains responsibility for validating the interface of those subsystems with its own components.

Delphi Technologies perform the majority of its ALT programmes on hydraulic test rigs, where the injected fuel is not combusted. However, some product failure modes are only excited by combustion, so additional engine tests are required. As such Delphi Technologies is provided with pre-production engines for internal testing and has visibility of the engine and pre-production vehicle testing completed by the customer. Delphi Technologies and its customers engineering teams then work collaboratively to investigate any product failures or performance degradation observed in engine testing.

A paper by Zielenski et al (2014) outlines the typical reliability requirements associated with Euro VI applications, the methods used in the design of reliability demonstration and ALT programmes, and more specific ‘targeted tests’ for individual failure modes implemented by Delphi Technologies. The paper highlighted that products for CV applications are typically developed in compressed NPD programmes, and that the introduction of common rail technology resulted in an increased uncertainty about product life in application than for previous products. The paper highlights the advances in methods and capabilities that were associated with the Euro VI NPD programmes, which continued with the authors role in the business at the time of writing this thesis.

5.4.3 Recording of failure incidents and problem-solving process

Product failures, including both failures associated with the NPD process and those associated with warranty claims, are logged by Delphi Technologies using a commercially available FRACAS tool. The failure modes associated with those failure incidents are then investigated using the 8D group problem solving process, which is in turn recorded through the same FRACAS tool. In Delphi Technologies, the 8D process is led by PV, but brings together a cross functional team to identify the root cause(s) of a failure, and to identify, implement, and verify corrective and preventative interventions.

In order to identify the root cause(s) of a failure mode, several RCA tools typical of the industry are employed by Delphi Technologies. The two principle tools used are FTA and 5Why, both augmented through empirical and analytical investigations. For more complex problems, Delphi Technologies utilise Shainin techniques coordinated through a centralised group of trained Shainin practitioners. It is also common for cross functional ‘taskforces’ to be identified for problems that span multiple product applications or fail to be solved through the previously discussed methods.

5.5 Knowledge Management & Expert Judgement in Delphi Technologies

Within the CV business unit of Delphi Technologies, the NPD process was historically governed by the 5P procedure as summarised in Figure 30. The 5P procedure represents a typical phasegated product development process, defining the NPD cycle in 5 phases, transitioning from development of the initial opportunity, through to mass production of the product. Each phase comprised several work packages owned by different functional groups within the business. The 5P procedure featured a total of 7 phasegates, where project status & forward plans were reviewed against both the customer requirements, and the internal project success criteria. The phasegates are identified by yellow diamonds in Figure 30. At each phasegate, the project could be either be approved to proceed, or rejected, with all evidence and actions captures by the Programme and Project Managers in the Project Evidence file.

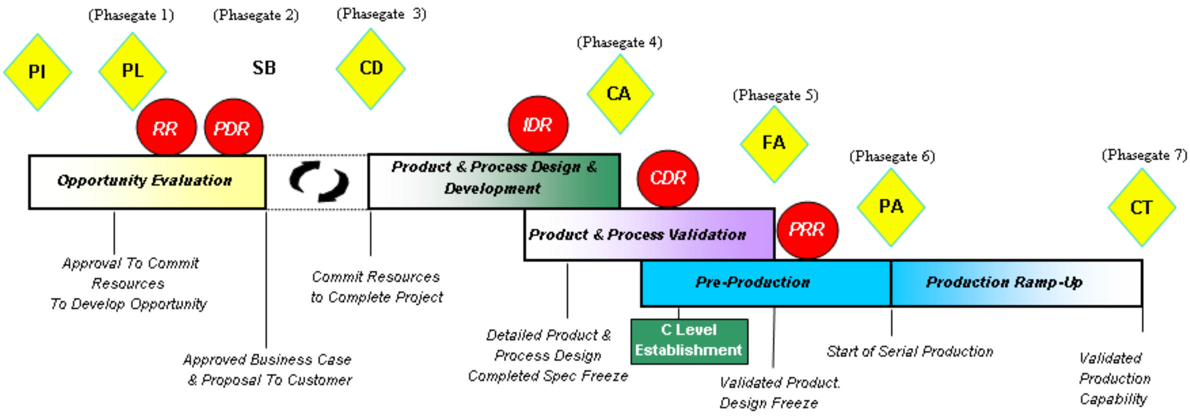


Figure 30: '5P' procedure for NPD

In addition to the management phasegates, the 5P procedure also featured 5 separate Design Reviews, the objective of which was to provide a forum to critique the technical aspects of the product design, process design, and product validation activities through the product development cycle, ending when the product enters pre-production as a fully validated product. The Design Reviews also act as a communication mechanism to escalate any technical risks associated with the product. At the end of each design review, the product engineering directorate then make an informed decision whether to proceed, hold, or re-design the product. By incorporating senior product engineering representatives from different functional groups, Design Reviews also served as a means to communicate technical concepts across projects.

In the Product and Process Validation phase of the 5P procedure, PV serves to identify failure modes which may affect product performance during its normal life, and also serves to demonstrate the products reliability to both Delphi Technologies and the customer alike. As such, progress against the

PV plan, and the key metrics of reliability growth, confidence growth, and the status of open failure modes are reviewed in both the Phasegates and Design Reviews.

The CV business unit of Delphi Technologies is split across three locations in the UK. One site (location 1) is a dedicated technical centre, featuring product engineering staff only, while another is a dedicated manufacturing site (location 2). The third site (location 3) has a mix of both manufacturing and product engineering functions. PV testing facilities are split across the two product engineering sites. For the Euro VI family of products, the business also split the engineering and manufacturing functions by components, with ‘pump’ and ‘injector’ components nominally being decoupled. Table 4 shows the results of this geographical, functional, and project-based split, with the location indicated for each element of the products highlighted.

Location: 1, 2, 3	Product A	Product B	Product C
Design	1	3	3
Development - Injector	3	1	1
Development - Pump	1	3	1
Validation Testing	1&3	1&3	1&3
Manufacture - Injector	2	2	3
Manufacture - Pump	3	3	3

Table 4: Geographical, functional, and project-based split for F2 NPD programmes

The CV business unit of Technologies adopts a matrix organisational structure for the product engineering functions. Project teams are constructed for each product with team members representing each function, often located across different physical locations. For a given location, project teams are generally collocated alongside members of other functional groups within the wider NPD process, such as Project Management and Manufacturing Engineers.

The Product Engineering element of Delphi Technologies can be shown to take the form of an Expert Driven Organisation (EDO) (Chang et al, 2011), where the contributions of several experts distributed across the business’ locations and functional groups, are taken as authority. EDOs are characterised as being strongest at knowledge internalisation (Chang et al, 2011), and this can be witnessed in Delphi Technologies. Often, the individual experts act as internal consultants, moving from across projects, applying their unique knowledge to solve problems as they arise, strengthening their own intellectual capabilities and experiences. Such a system makes use of the experts in solving the most difficult problems facing the business, but when combined with a reduced emphasis on sharing that knowledge through Knowledge Management tools, fails to promote wider knowledge combination within the organisation. Outside of their allocated project roles, the technical experts within Delphi Technologies can be problematic to access. Distributed across several sites in the UK, with busy diaries and customer driven deadlines, one-on-one expert elicitation can prove difficult while group engagement, such as a workshop, can prove next to unobtainable.

During the development of the Euro VI family of products, Elicitation of expert judgement in the Delphi Technologies NPD process was both informal and widespread. Figure 31 shows the numerous points in one section of the highest level of the Delphi Technologies PV procedure where expert judgement is utilised, highlighted in red, in addition to objective data such as calculations, models, & statistics. These points took many forms, from formal group FMEA workshops, where expert opinion on the severity and likely frequency of product failure modes is elicited, to the application of expert judgement in assigning a pass/fail decision to optical inspection of post –test FIS.

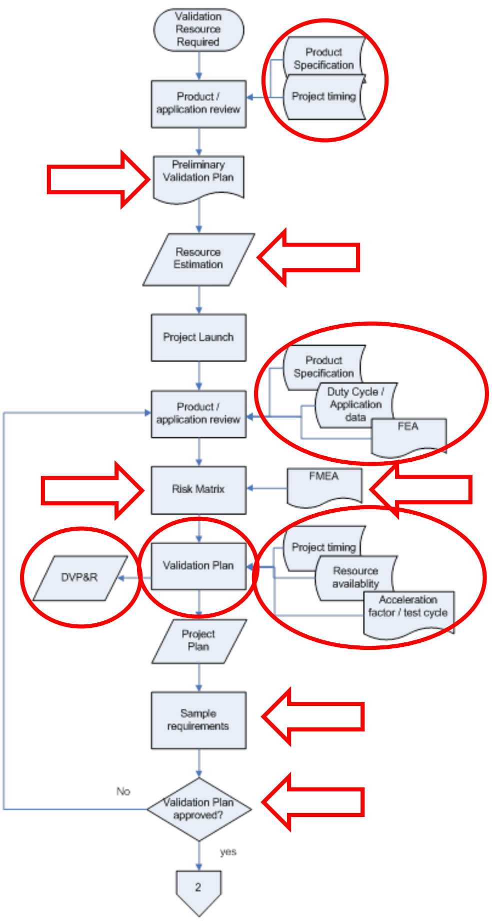


Figure 31: Part of the high-level PV procedure with expert judgement utilisation highlighted

5.6 Experimental design in Delphi Technologies

Within the CV engineering functions of Delphi Technologies, formal use of experimental design methodologies has historically been limited to certain functions or specific use cases. As has been demonstrated as typical of the automotive industry in §4.4.6, its use within Delphi Technologies is prevalent in optimisation of engine calibration and in analytical experiments using CAE tools. However, in the context of failure mode characterisation, empirical test programmes are typically designed using either OFAT, Best Guess, or multiple factor at a time methods. Often this is as a result of physical resource constraints rather than intellectual constraints, where neither sufficient samples or test facilities are available, or when sufficient time is not available for the planning, execution, and analysis of more formal Experimental Design methods.

5.7 Summary

This Chapter has provided an overview of the sociotechnical context of this research. The process for injector performance characterisation at Delphi Technologies has been introduced, identifying changes in the fuel gain curve as a key performance indicator. The application of Knowledge Management in the NPD procedure was discussed, and Expert Elicitation was demonstrated to be widespread within that procedure, but often informal in its application. The Product Validation process has also been discussed, showing how ALT and DFR tools are used in a lean engineering process, typical of the industry. Finally, it was shown that experimental design is not frequently used within the organisation, as typical of the industry, and typically not applied in FMC.

The next Chapter will introduce the case study of FMC associated with this Thesis, discussing how Expert Judgement was Elicited through a structured Delphi Study.

Chapter 6 NCV Seat Wear – Eliciting and Codifying Expert Judgement

6.1 Introduction

This Chapter introduces the first element of the case study of the FMC process outlined in this thesis. Structured Expert Elicitation is presented as a key element of a systematic approach for exploring uncertainty in complex systems in order to better structure future investigation. The Delphi Method will be presented as a flexible technique for eliciting expert judgement in such complex systems, and the design of such a study using a cross section of the experts within Delphi Technologies will be presented. The results of that study, in the form of a group led definition and ranked taxonomy of variables that influence the failure mode is presented. Causal Loop Diagrams will then be used to build a system model to describe that failure mode, capturing interactions between variables, and the level of consensus around significance of different factors. A visualisation of this Chapter shown in Figure 32.

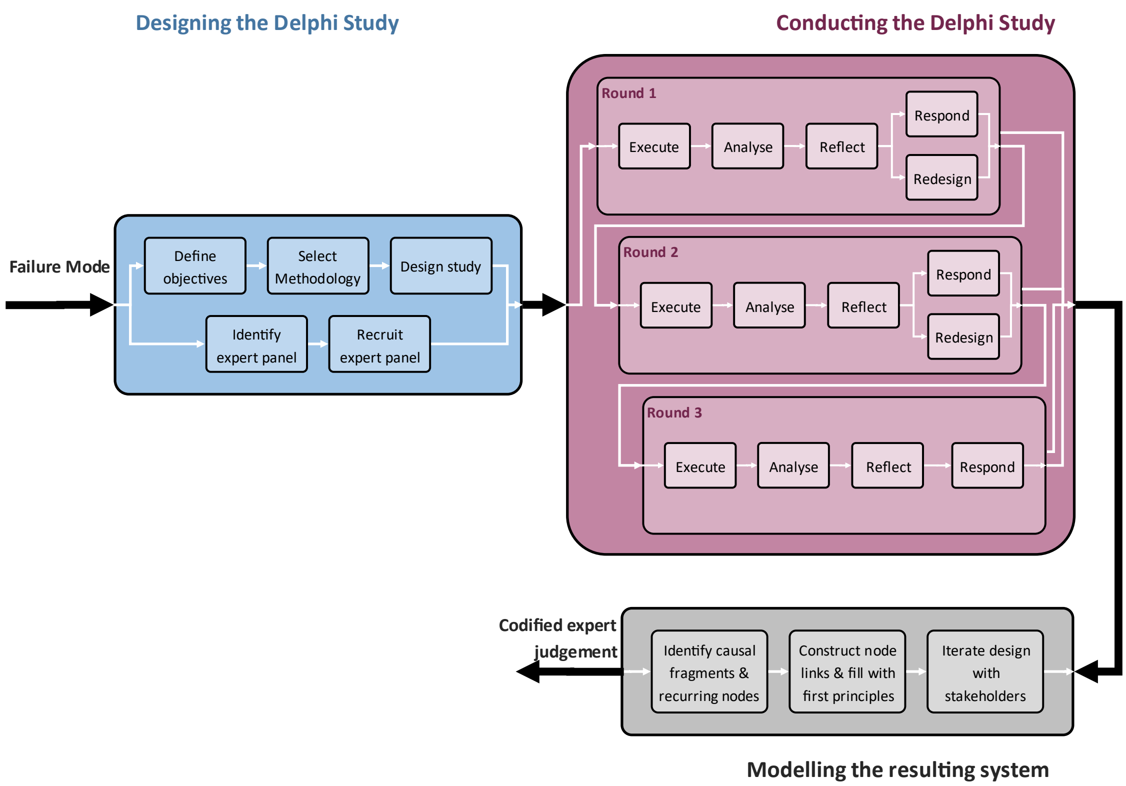


Figure 32: System diagram for this Chapter

6.2 Introducing the case study and the need for Expert Elicitation

In this case study, the subject is a failure mode that has been present for a number of years, across different products and applications, ultimately resulting in an observed fuelling increase in the ballistic region of the gain curve. As discussed in Chapter 3, as emissions limits have become more stringent, sensitivity to changes in fuelling around the quantities typically used for pilot and post injections has increased, so the perception of the significance of this failure mode has increased in parallel.

Having identified possible root causes through FMEA and RCA, an indicative example of which is shown in Figure 33 in the form of a top-down, deductive hierarchical failure model, previous empirical and analytical work had partially or completely ruled out a significant number of potential root causes. This research then focuses on wear to the bottom seat of the 3-way NCV, herein referred to as NCV bottom seat wear. The corresponding interface is highlighted in Figure 34. NCV bottom seat wear can result in an increase to the effective lift of the valve, and a change to the flow characteristics across the valve, both of which are potential root causes for the fuelling change. As valve lift is low by design (<<100um), to increase response time and minimise leakage, even a few microns of material removal can result in a significant % change.

Figure 33: Hierarchical failure model for injector fuelling drift (deliberately illegible), where Green boxes indicate 'Eliminated from Investigation'

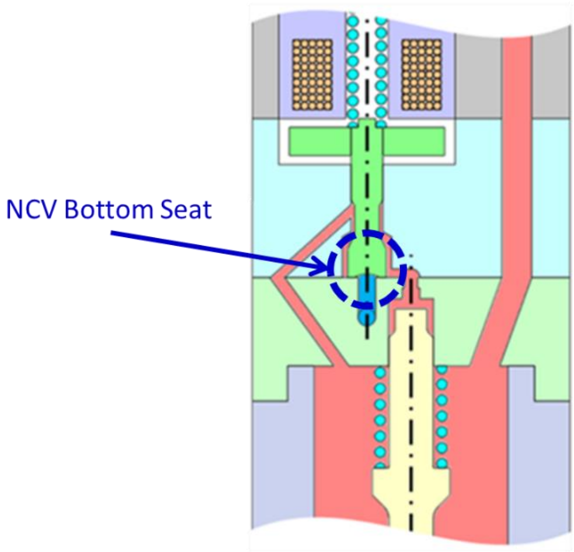


Figure 34: NCV bottom seat interface

Despite numerous parallel and tangential investigations, through the 8D process, Shainin investigations, and dedicated taskforces, the failure mode had yet to be fully characterised, meaning that while fragments of knowledge about the failure mode existed, there was uncertainty of the detailed failure mechanism, the specific variables that influenced it, and rate with which it progressed. This partial characterisation prevented both robust design solutions from being identified, and a fully representative and repeatable PV test from being defined such that those design solutions could be assessed. This represents significant uncertainty in the NPD process, to the detriment of the efficiency through which new products can be brought to market while retaining a flawless launch.

With this failure mode having been present over numerous different product platforms over a significant period, there was potential for significant levels of uncodified knowledge distributed through the business that could potentially inform the design of a formal, and comprehensive empirical study. The experts who held this knowledge were distributed geographically, managed through a matrix organisational structure, and engaged in parallel NPD programmes, representing a complex sociotechnical system. Indeed, such was the complexity it was perceived by the stakeholders of this research that you could ask any given expert on any given day of the week, and you would get different root causes proposed.

Given the numerous previous empirical and analytical investigations that have focused on this failure mode in the past, it was anticipated that this study would perhaps not generate any significant new knowledge around specific elements that influenced the failure mode. Instead, it was hoped that the study would combine that knowledge into a single, comprehensive whole, in a systematic way, such as to best structure further investigations.

The first element of this case study therefore had the objective of transforming that distributed tacit knowledge, into codified knowledge in the form of a group led definition of the failure mode, and a cohesive system model that described the variables that influence progression of the failure mode, and how they interact. The case study also seeks to establish consensus, or informed disagreement on which of those variables are of the most significance to NCV bottom seat wear, and therefore ballistic fuelling drift. There was also a specific desire to establish which usage variables result in acceleration of seat wear such as to inform future empirical investigations and ultimately ALT.

As presented in §3.7, wear to such valve seats can be classified as either cavitation, hard particle damage, impact wear, or sliding wear. In the experience of Delphi Technologies, it was possible to limit the effects of cavitation wear through optimisation of material coatings and hydraulic design analytically, as verified through empirical results. Similarly, it was perceived that debris damage could typically be controlled through material coatings, fuel specifications, and cleanliness standards for manufacture. As such, it was generally perceived that the wear to the NCV bottom seat, and the subsequent fuelling drift, would be associated with either impact wear, sliding wear, or a combination thereof. There was also a perception in Delphi Technologies and its customers that the problem was typically worse on engine or vehicle, with multiple hypotheses as to why, but that perception was unconfirmed. That perception was visualised by Lloyd (2015) through the working mental modal of the problem, expressed as ballistic fuelling drift, as shown in Figure 35, where 'DDS Rig' refers to unfired hydraulic test rigs.

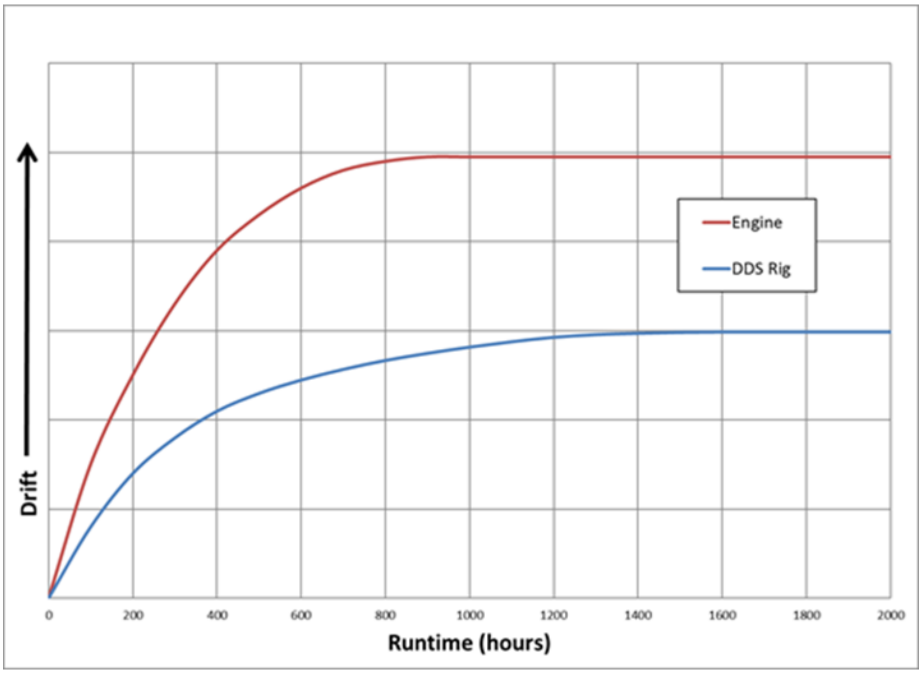


Figure 35: Working mental model of injector performance drift over time

This working mental model is notable in being asymptotic, which is atypical of most models of physical wear. However, this asymptotic behaviour can be observed in experimental results with significant sample sizes, and several assumptions are proposed that serve to support the validity of the mental model. Firstly, this mental model is presented in terms of injector fuelling drift, and while NCV bottom seat wear was proposed to be a significant contributor to injector fuelling drift, there are other, potentially confounding failures modes that progress differently over time, including those that are only present on engines, such as nozzle coking as a result of combustion deposits, that would negatively influence injector fuelling drift. Furthermore, as the valve wears, its geometry is likely to change, influencing the contact pressure profile at the component interface, and thus wear rate. Finally, as both surfaces of the valve interface feature multi-layered DLC coatings, sufficient material removal will result in a different combination of coating layers in contact, influencing the relative hardnesses of the interface. When characterising diesel inlet valve wear, Lewis & Dwyer-Joyce (2007) presented empirical results for valve wear that demonstrated progression similar to that of this mental model in some treatment combinations, then codified in their models.

6.3 Eliciting Expert Judgement through the Delphi Method

6.3.1 Motivations for use of the Delphi Method

The literature reviewed in §4.2 has shown that the Delphi Method offers a flexible technique for exploring concepts both within, and outside of, the existing body of knowledge with a group of experts. The method also allows for a flexible design with respect to both innovation and practicality, allowing the user to balance application considerations with validity. Other factors influence the validity of an application of the Delphi Method have been identified by literature, associated with both methodological weaknesses and more specific deficiencies in application. However, those same limitations on validity are not unique to the Delphi Method, but instead are true of other qualitative group elicitation strategies.

The technical experts within Delphi Technologies most suited for this study comprised of a number of individuals split across different geographical locations, supporting different projects, and occupying a range of roles within the organisation. Furthermore, many exhibited strong opinions and personalities, and previously, their opinions had not been very well conveyed, and/or received by their peers and superiors. The anonymity and avoidance of direct face-to-face interaction afforded by the Delphi Method was therefore chosen to allow those experts to contribute to the study in the most effective manner.

The availability to attend physical meetings of geographically distributed technical experts with customer focused schedules can be greatly restricted. Indeed, in this study, the first available meeting opportunity with the entire expert group would have been limited to a 1-hour session, 5 weeks from the time of scheduling, while a more suitable half-day, co-located workshop would not have been possible to schedule within 6 months. In the context of this study, typical of many firms engaged in NPD, the distributed, remote engagement of the Delphi Method would allow for several iterative rounds of exploration to be completed before even the most restrictive of physical meetings could have been completed in this environment. The Delphi Method was therefore chosen as a resource sensitive approach to Expert Elicitation.

6.3.2 Alternative methodological choices

With existing investigations having taken place around the business, it may have been possible to aggregate the codified results of previous empirical investigations, before then summarising what conclusions and uncertainties existed, informing future empirical investigation. However, such a method would potential exclude the tacit knowledge held within the expert community of Delphi Technologies, being reliant on that which is captured in reports, test results, and other communications. Furthermore, this approach was not deemed suitable by all stakeholders, expressing concerns that it would be possible to simply repeat, or retry previous investigations. A more structured approach, that captured not only the existing codified knowledge, but the experience of the expert community such that lessons learnt could be incorporated, was sought.

Some other forms of Expert Elicitation, including the Nominal Group Technique (NGT), are presented as being suitable for combining the judgement of a group of experts where confrontation may exist. While NGT can promote equal participation of all group members, its more rigid structure does not necessarily lend itself to be flexible in handling emergent themes, or combination of knowledge, a flexibility afforded by the iterative nature of the Delphi Method. Furthermore, in requiring the expert panel to attend a co-located workshop, the NGT technique would not be easily accommodated in Delphi Technologies, with the distributed, and time-bound nature of its expert community.

6.3.3 Objectives of the study

This study had 3 main objectives. Firstly, group consensus on a definition of the problem was sought as a boundary for further investigation. Secondly, a detailed characterisation of the failure mode was sought through exploration of its influencing variables and their interactions in order to inform further empirical and analytical investigation. Finally, the panel’s worldviews, and their perceptions of the reasons for why Delphi Technologies had previously been unable to fully characterise the failure mode, were sought.

6.3.4 Design of the Study

The application of the Delphi Method for this study has been designed with a view to the criticisms and lessons presented previously. A qualitative study was chosen to facilitate exploration and theory building around the characterisation of the failure mode. The study was conducted remotely via email, blindly distributed to the panel, with files attached by both the author/facilitator and the panel members alike. As several panel members shared the same open plan office space as each other and the author/facilitator, the panel were explicitly asked to not discuss the study directly with their colleagues, or the author/facilitator, in order to retain the desired anonymity.

For this study, while the design of the Delphi Method was conceived with a notional framework, it was recognised that some evolution in design may be desirable. The application was designed to have 3 rounds, based on the 3 main objectives for the study, with the opportunity for the panel to review and reflect upon the group responses of the previous round. The panel members would then be able to iterate their responses in light of those group responses if their own judgement was influenced. The design intention was that the panellists would be able to complete each round in 30 minutes, but still allow sufficient scope for rich feedback as appropriate.

The first round of the study was designed as an opening round, with 3 short questions intended to promote engagement with the panel. The panellists were to be asked to provide their definition of the failure mode, identify any variables and their interactions that they judged to be of influence, and to establish their worldviews on the plausibility and desirability for the product to be robust to the failure mode, or whether they thought there were alternative failure modes that should instead be prioritised.

The design intention of the second round was to iterate the response to the first round in light of the group's collective response, before then exploring the relative perceived significance of the variables by asking the panel to weight each on a defined scale. The design intention of the third round was to once more to iterate the results of the previous round, before then asking the panellists to suggest the possible causes that have previously prevented Delphi Technologies from fully characterising this failure mode.

Twelve experts were invited to take part in this study, selected in conjunction with key research stakeholders based on their expertise in the subject matter and their availability and motivation to take part in such a study alike, forming a homogeneous panel that met the minimum size recommendations identified in the literature. Subject matter expertise was defined as having direct experience in leading, or participating in, investigations into ballistic fuelling drift across any product platform past or present. A number of the panellists were no longer directly involved in such investigations, or were no longer part of active product development functions, but were still identified has having relevant previous experience. No panel member had any prior experience with the Delphi Method, but all had participated in other forms of Expert Judgement Elicitation such as FMEA or 8D problem solving. The expert panel is summarised in Table 5, with their current functional groups, and relative positions and experience in the business identified, representing over 140 years of combined experience. Figure 36 then visualises the relative position of each across the functional groups and hierarchy of Delphi Technologies, with both the researcher and the research stakeholders identified.

	Current Functional Group	Seniority and experience in the business
Expert A	Development	Team Leader, 10+ years
Expert B	Development	Manager, 10+ years
Expert C	Validation	Principal Engineer, 10+ years
Expert D	NPI Manuf. Support	Principal Engineer, 20+ years
Expert E	Analysis	Manager, 10+ Years
Expert F	Test Rig Support	Manager, 10+ years
Expert G	Development	Prinicipal Engineer, ~10years
Expert H	NPI Manuf. Support	Team Leader, 20+ years
Expert I	Design	Principal Engineer, 15+ years
Expert J	Development	Senior Engineer, 15+ years
Expert K	Validation	Senior Engineer, 5+ years
Expert L	Development	Senior Engineer, ~10 years

Table 5: Expert panel members

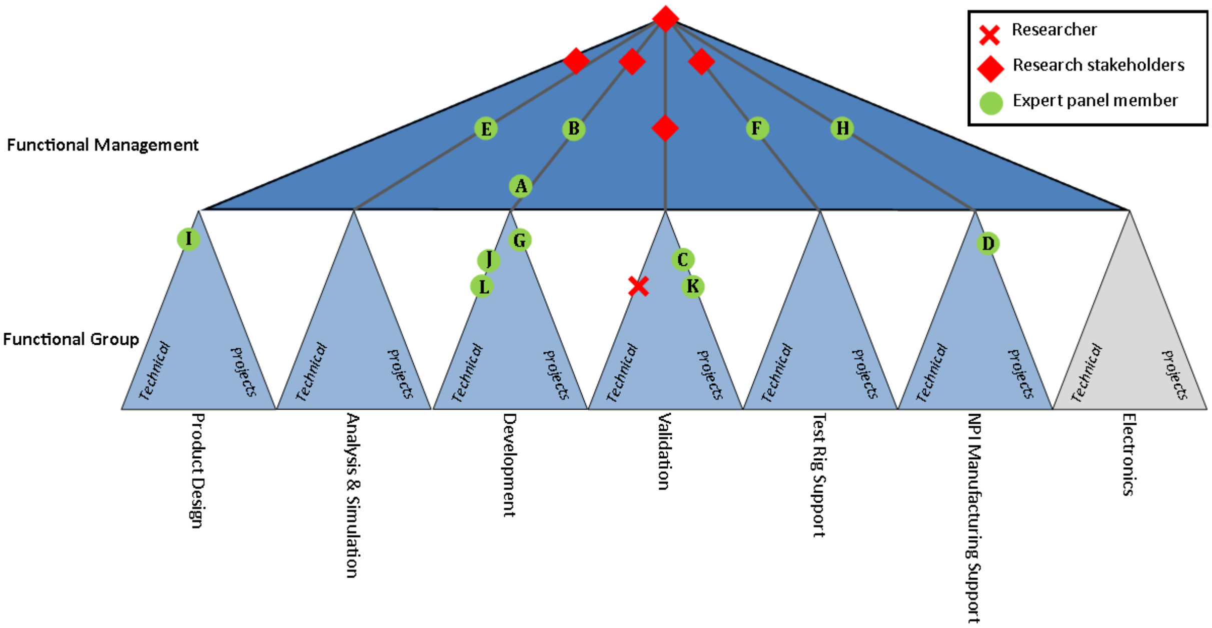


Figure 36: Relative position of each expert within the functional groups and hierarchy of Delphi Technologies

To facilitate both initial engagement and panel retention, several strategies previously identified in the literature were implemented for this study. As part of the invitation to participate, all panel members were provided with an information pack, detailing: the relevant background and motivation for the study; an overview of the Delphi Method; a timeline for completion of the study, including the dates associated with each round; and a ‘terms of participation’ document, outlining the expectations for the panel and the facilitator alike. The invitation took the form of an email providing a concise overview along the questions for round 1, with the information pack attached to prevent the panel being overloaded with material, allowing them to best consider the invitation to participate. The subject title of the email was chosen carefully to stand out from business as usual, and the logo of the supporting academic institution featured on all material supplied to the panel to enhance the creditability of the study. Through the course of the study, the panellists were at all times referred to as 'experts', with the exception of times when more personalised communications were appropriate.

In order to limit the time and effort required to complete the study, effort was made to limit the scope of each round, and to provide template for responses, such that each round could be completed in 30 minutes or less, as outlined in the agreed terms of participation. To best ensure that the study adhered to the agreed timeline of the programme, panel members were reminded before the close of each round, with a 1-day window for late submissions to their availability. The results of the previous round, along with any new questions, were then provided within 2 working days, adhering to the terms of participation whilst reflecting the panel retention strategy.

The final design of the study is as summarised in Table 7.

Methology	Qualitative
Expertise Criteria	Experience, relevance, motivation & availability
Purpose	Exploration and theory building
Number of participants	12; homogeneous sample
Number of rounds	3 x ~30minutes
Anonymity	Full requested
Mode	Remote via email and attached files
Rigour	Results and decisions auditable. Researcher concious of own bias resulting from closeness to subject
Results	Consensus indicated by an adapted 'Italian Flag' notation Qualitative justifications provided uncensored
Verification	Identification of future experimental studies & related research questions

Table 6: Summary table of Delphi study design

6.4 Results of the Expert Elicitation Study

6.4.1 Efficacy of the Study

The 3-round structure of the study as outlined in the terms of participation was adhered to, but the specific structure of each round required some evolution in order to derive best value from the study. While rounds 1 & 2 remained largely as designed, a decision was made to alter the focus for round 3 as a result of a combination of emergent factors. Firstly, the ranking of variables completed in round 2 showed significant disagreement between the expert panellists. Secondly, the panel had exhibited a consistently low willingness to reflect upon, and iterate the results of previous rounds, instead focusing on progressing with any new questions, to the detriment of the iterative nature of the Delphi Method. This is either indicative of a flaw in the methodology or the application, or perhaps a symptom of the environment the expert panel are used to working in.

In addition, several members of the panel did not provide justification for to their rankings provided in round 2, potentially making it difficult for the remainder of the group to reflect with respect to their own judgement. As such, round 3 was re-designed to explicitly task the expert panel to reflect upon and re-rank the disputed variables identified in round 2 in light of the justifications provided that were by the panel. This lack of justification could indicate that either the panel members were unsure of their opinions, or that they were time-bound in their response. The question that was originally designed to be asked in round 3 was instead asked as an optional closing question. This provided the panel an option to participate in a fourth round without deviating from the agreed terms of participation.

The panel's responses were tracked through the study to allow reflection on the application of the Delphi Method. Figure 37 shows the responses received from each panel member for each round, including the optional round 4. No nominations of additional panellists were received as a result of the initial invitation, and one panel member chose to opt-out from the study citing a perceived lack of knowledge of the failure mode in application. A further expert then chose to opt-out after round 1, citing the results presented as being inappropriately 'too vague or too specific'. As can be seen, a combination of reasons, including annual leave and specific customer deadlines, led to varying participation levels within the panel, providing evidence of the limitations of EE in resource sensitive environments.

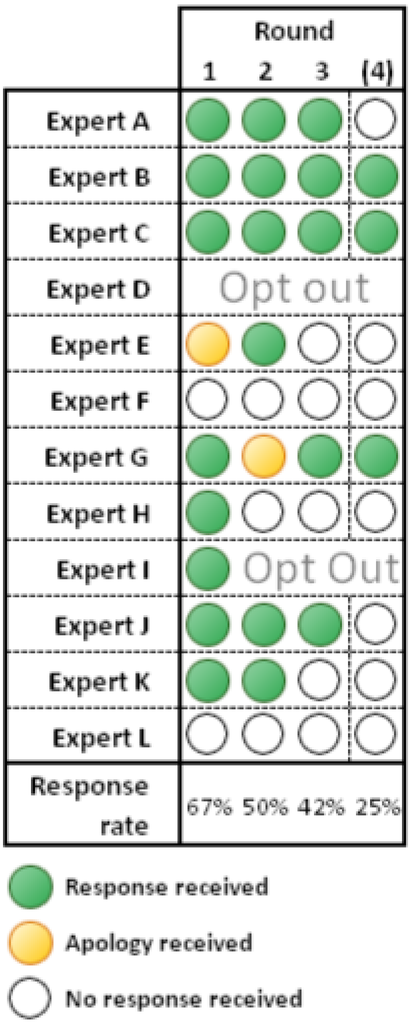


Figure 37: Responses for each panel member per round of the study

Figure 38 shows the response rate for each round, which can be seen to decline in an approximately linear fashion. This observation could be indicative of either panel fatigue, or the more general issues associated with the availability of experts within Delphi Technologies.

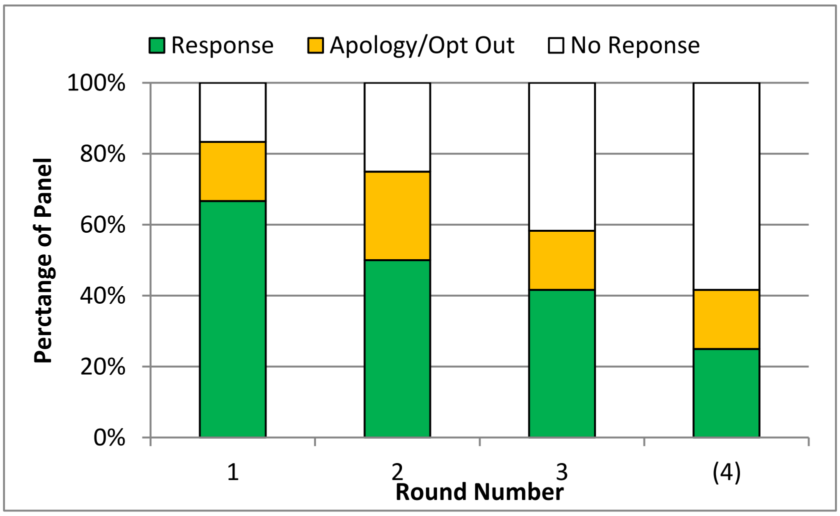


Figure 38: Response rate per round of the study

Figure 39 then shows the cumulative responses for each working day the study was open for. As can be seen, by the close of the first working day, 5 responses had been received demonstrating that the study generated high interest and engagement in the experts within Delphi Technologies.

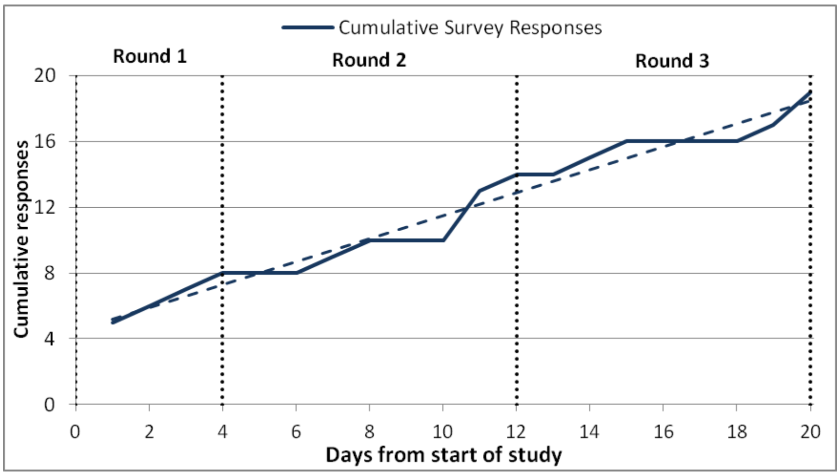


Figure 39: Cumulative responses from beginning of the study

The anonymity of the study was partially breached through discussions of the study in an open-plan office space, and through some experts suggesting they could identify fellow panellists through their responses, highlighting a potential limitation of an internal, homogenous panel.

An expert elicitation study utilising the Delphi Method through a 3 round, electronic and distributed interaction, has been selected as a method to explore the failure mode in this case study. This method was selected due to the remote and anonymous interaction with the expert panel it afforded. The study was designed for this specific application with reference to the lessons learnt from the literature. While the initial response to the study was high, the subsequent rounds demonstrated the limitations associated with engaging experts in busy, customer centric roles.

6.4.2 Group Led Definition

In round 1 of the study, the expert panel were asked to define the failure mode. The individual panel members were asked to provide a free text response that would then be transformed into an aggregated response through qualitative data analysis. The author/facilitator assessed the responses, capturing any recurring themes and identifying the emerging consensus. The composite result included sentence fragments provided directly by the panel when appropriate. The proposed definition was then presented to the panel:

NCV Seat Wear: *"A physical change, be it a removal or displacement of material, through sliding wear, at any of the 4 seating surfaces that make running contact resulting in a change in seat geometries. This change in geometry influences the effective NCV lift, to the detriment of fuelling stability over life"*

In round 2, the experts were asked to review this proposed definition, and either approve or suggest changes as appropriate. No changes were suggested by the panel and the definition was approved.

With respect to the discussion of NCV seat wear presented in §3.7.3, this definition is notable in referencing only sliding wear, with no inclusion of wear associated with impact or flow erosion.

6.4.3 Variables and interactions

When asked in round ` to identify the variables that influenced Control Valve seat wear, the panellists identified a total of 31 variables, grouped by them as either pertaining to the design, manufacture, or usage of the Control Valve. The identification of unique variables required an element of judgement from the author/facilitator to identify instances where different terminology was used by panellists, highlighting the requirement for an agreed taxonomy and shared language. The grouped variables were presented in the feedback to the panel, and no comments, clarifications, or changes were suggested by the panel.

In round 2, the experts were asked to explore the interactions between the variables that were identified in round 1. Markedly different levels of resolution were employed by the experts during this process, with some focusing only on the highest-level interactions, listing 10 interactions, while others identified every action regardless of perceived significance, listing over 100 interactions.

The reliability of the experts was assessed by their consistency in correctly labelling pairs of interacting variables, in a similar manner to the 'paired comparisons' metric employed in literature. Only one expert (B) made any mistakes, failing to correctly label 15 paired interactions, but given that they identified significantly more interactions than any other expert, this only resulted in a 7% error rate.

The group results were aggregated using matrix addition for the variable pairs identified by each expert. The resultant matrix was such that any the higher the value for each variable pair, the greater the level of consensus of its significance to Control Valve seat wear. This was summarised for the feedback to the panel using a colour gradient, an example of which is shown in Figure 40.

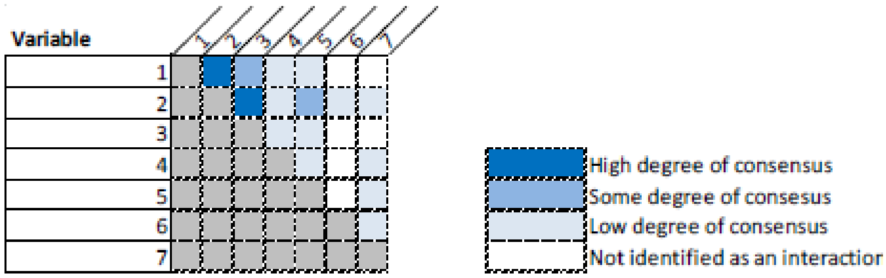


Figure 40: Example of feedback for variable interaction consensus

In round 2, the panel was also asked to assess the significance of the variables identified in round 1, reviewing review the variables, along with any justification or evidence supplied by their peers, and rank each one as being of either high or low significance to Control Valve seat wear, to be not applicable, or to withhold their judgement. While some experts assessed the significance of each variable, others elected to only rank those that they considered as either of high significance, or as not applicable, suggesting that they chose to withhold their judgement on the variables they were less certain of, or that they only wished to save time in responding by only ranking the most and least applicable variables. A total of 4 new variables were also identified by panellists in round 2 and were presented to the panel for their further review.

In addition, the experts were asked to provide justification for their judgements to share with the panel. However, many experts provided no justification or evidence to support their rankings at this time.

The results of this ranking process were presented back to the panel using an adapted 'Italian Flag' notation (Blockley & Godfrey, 2000), identifying the degree of consensus for each variable as shown in Figure 41. The author/facilitator then pooled the variables into 4 groups: those with a consensus as being of high significance; those with a consensus as being of little to no significance; new variables identified in round 2; and those were contradicting judgements were witnessed. Despite one expert electing to opt out after round 1 as the variables identified were either too specific, or too vague, no refinements to the definitions, nor suggestions for either splitting or merging variables, were received.

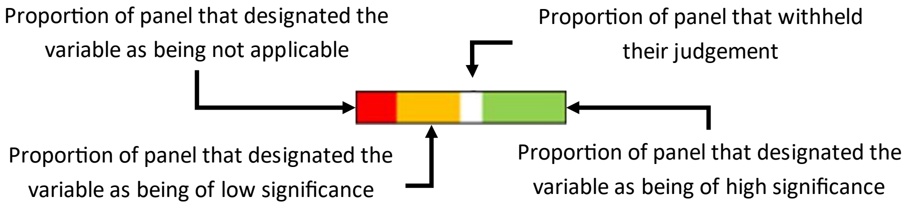


Figure 41: Example of feedback for variable consensus as significant

In round 3, the panel was asked to review the rankings and evidence for the 17 round 2 variables where no consensus was met, and the 4 new variables identified in round 2, and iterate their own judgements of their significance. The panel was supplied with a summary of the anonymised, verbatim rankings and justifications provided by their peers in round 2, along with the interactions identified for reference. The panel was asked to reflect on this summary and provide a second ranking with revisions and justifications as appropriate. The proportion of experts that provided qualitative justifications alongside their rankings increased in round 3 although some experts still failed to provide any comments alongside their decisions.

When presented with the evidence and justifications supplied by their peers, the majority of the active panel members chose to revise some of their own judgements, with the exception of one panellist who elected to make no changes to his previous judgements, nor did he provide any justifications in doing so. Of those experts who did elected to change their own judgements: one acknowledged that he lacked evidence to support their previous judgement; one agreed that some variables were open to differing interpretations; and others provided no justifications at all.

The results of this iteration were summarised through a visualisation of the relative consensus before and after round three, alongside a comment from the author/facilitator on the degree of any change in consensus, an example of which is shown in Figure 42. The majority of the 17 previously discordant variables showed a degree of change in relative consensus after round 3, but only 4 moved to consensus as either being of high or little to no significance to the failure mode. No consensus was reached on the 4 new variables identified in round 2, suggesting further exploration may be appropriate as completed for the original variables.

Variable	Round 2 judgement	Round 3 judgement	Summary of Round 3
Operating pressure			Towards consensus as significant

Figure 42: Example of feedback of change in consensus between rounds

One variable, relating to the surface finish of the relevant features as manufactured, almost reached consensus as being of high significance. While the majority of the panel suggested it was of high significance, one expert was of the opinion that it was of no relevance, suggesting the as manufactured surface finish became irrelevant after a short bedding in period in operation.

Three variables almost reached consensus as being of little to no significance to NCV seat wear. Given the justifications provided by experts on both sides of the argument, there is possible evidence of significance being placed upon historic events that may have been related, but where no conclusive evidence could be supplied.

After round 3, 13 variables, including the variables added in round 2, showed no consensus in the judgement of their significance amongst the expert panel. Given that the variables had been reviewed by the panel, with evidence and justifications provided by the experts, with no consensus being met, these variables can be described as having reached informed disagreement on their significance.

Throughout this process, only one expert referred their judgements on variables back to the group-led definition which referenced only sliding wear of the seating components, with no inclusion of either impact wear or flow erosion. However, several variables identified by the panel were either solely, or significantly related to impact or flow erosion rather than sliding wear. One expert commented as such through the study, identifying the variables as not being of significance to the failure mode as defined. This suggested that the group-led definition perhaps required further iteration as the panel had identified variables deemed of significance that were not associated with the wear mode it described.

A final summary of the 35 variables, including the original variables and the 4 identified in round 2, was presented, with the variables grouped as:

- Those with consensus as being highly significant – 9
- Those almost at consensus as being highly significant – 2
- Those where informed disagreement was reached – 13
- Those almost at consensus as being of little to no significance – 3
- Those with consensus as being of little to no significance – 8

In presenting the results of this study, the panel were provided with a summary of the degree of consensus met on each variable, as shown in Figure 43. In addition, for each variable where informed disagreement was reached, a summary of all evidence and justifications was collated as an output of the study.

Variable	Expert judgement	Overall Consensus
Material specifications (inc. coatings)		General Consensus - Variables of high relevance to NCV seat wear
Seat geometries (diff. angle and contact area)		
Number of injections		
Fuel lubricity		
Geometric tolerances		
Deburr		
Operating pressure		
Hard particle debris contamination		
Tilt		
Surface finishes (Design)		Near Consensus - Some contrary opinions amongst experts
Surface finishes (Manufacture)		
NCV pin dynamic motion		Informed disagreement - Process has not yieled consensus
Fuel viscosity		
Operating temperatures		
Pin stiffness		
Flow area/rates across seats		
Assembly of pin & armature		
Pin mass		
Oil-in-fuel dilution		
Injection durations		
Pressure variation (dilation)		
NCV lift		
Concentricity tolerance		
Air ingress		
Low pressure system backpressure		Near Consensus - Some contrary opinions amongst experts
Wash & cleanliness		
Fuel Initial Boiling Point		
Clamp load distortion		General Consensus - Variables of low-to-no relevance for NCV seat wear
Tool wear		
Seat profile waviness		
Biofuel content		
RDO positional tolerance		
Fuel deposits		
Water-in fuel dilution		
Biofuel content degradation		

Figure 43: Summary of degree of consensus for each variable after completion of the study

6.4.4 Worldviews on the feasibility and desirability of robustness against seat wear

In order to capture the world views of the expert panel members, in round 1 they were asked to comment on their perception of the feasibility and plausibility of Delphi Technologies' products being robust against Control Valve seat wear. The panel's responses were consistent in their response that it is desirable for Delphi Technologies' products to be robust against control valve seat wear, but the group response suggested that it is not plausible to be 100% robust against seat wear given the current design concept. Instead, the group suggested that engineering effort should focus on reducing it to the point of insignificance with respect to the life of the product through the identification of the root cause(s) and implementation of appropriate design solutions. The group also agreed that an alternative solution would be to reduce the products sensitivity to control valve seat wear, even if to the detriment of as-new performance.

Despite the panel members representing a wide cross section of functional groups, levels of seniority, and physical locations, the world views of the panel on the desirability and plausibility of improving robustness to Control Valve seat wear were consistent, with no disagreements identified.

The panel were also asked to identify any through-life failure modes that they identified as being of higher priority to the business. Panel members identified a total of 4 wear-out failure modes judged to be significant, 2 of which being identified by multiple experts, but agreed that none were of higher priority than Control Valve seat wear at the time of the study.

6.4.5 Soft System causality

As the study was redesigned to encourage further iteration of judgement in the characterisation of the failure, the panellists were provided with the option to participate in a fourth round to explore the final objective of the study. Three panellists opted to take part, providing free text responses for what they judged as the reasons why Delphi Technologies had been previously unable to fully characterise this failure mode. While participation in this round was low, the quality and quantity of the qualitative feedback was markedly improved, Furthermore, one of the participants for round 4 had previously provided little to no justifications for their responses in previous rounds, suggesting that this question, and the opportunity to explore the subject, motivated those that did participate into sharing more of their subjective thoughts. The results of this analysis will be discussed in §6.5.3.

6.5 Generating system models using the results of the Delphi Study

This section will present the process through which the results of the Delphi Study, both in the form of the exploration of the failure mode and the soft system causality behind the previous inability to full characterise the failure mode, were transformed into System Models using Causal Loop Diagrams.

6.5.1 Failure mode causality model

The Delphi Method applied in §6.4 resulted in the identification of 35 variables, with interactions between them, and resultant failure explicitly discussed by the expert panel. The author applied qualitative data analysis to identify ‘causal fragments’ associated with the variables, linking the concepts identified in the expert’s judgements. For example, if the expert panel identified that a change in the ‘fuel return back pressure’ would influence the ‘damping’ acting on the valve, changing the ‘NCV pin impact velocity’, then a causal fragment could be identified, linking these three variables in series. This process was completed by constructing the fragments on a large sheet of paper for flexibility and simplicity but could be similarly achieved using a variety of software tools. These causal fragments took the form of multiple node and link clusters, and example of which is shown in Figure 44, where the aforementioned example is found in the top left.

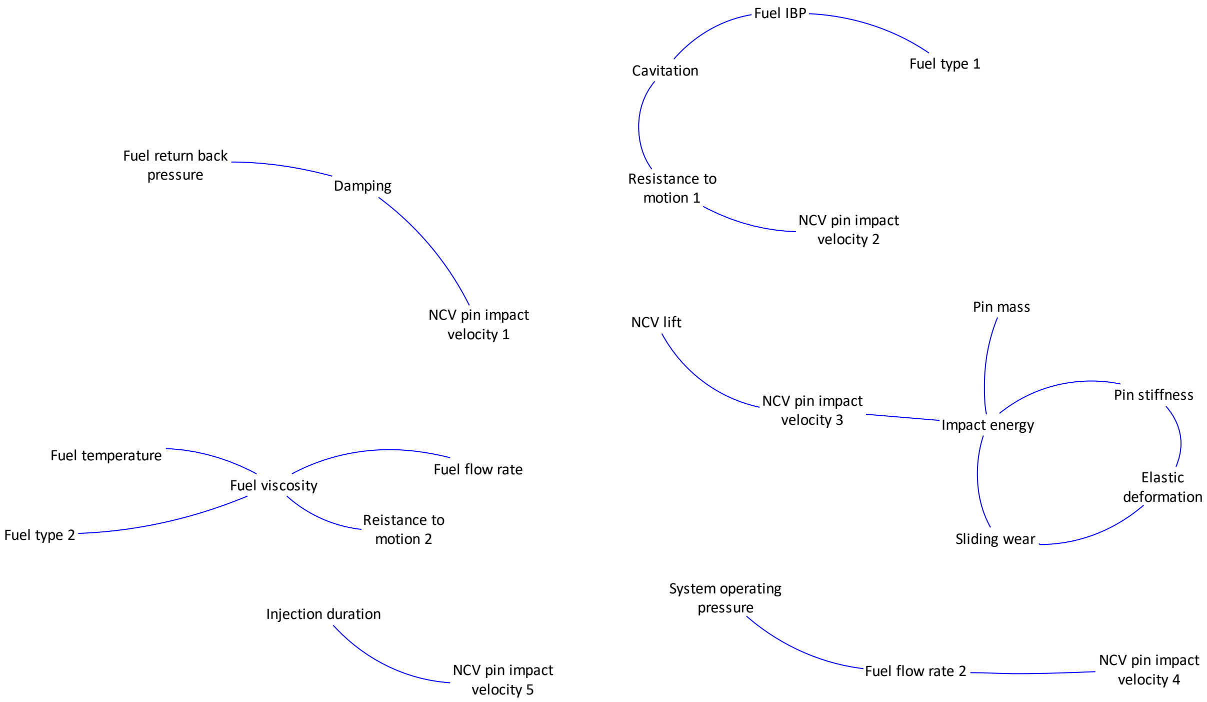


Figure 44: Example of causal fragments

Recurring nodes were then identified through colour coding to facilitate model building, an example of which is shown in Figure 45, where 'NCV Pin impact velocity', 'Fuel type', 'Fuel flow rate', and 'Resistance to motion' are all recurring nodes.

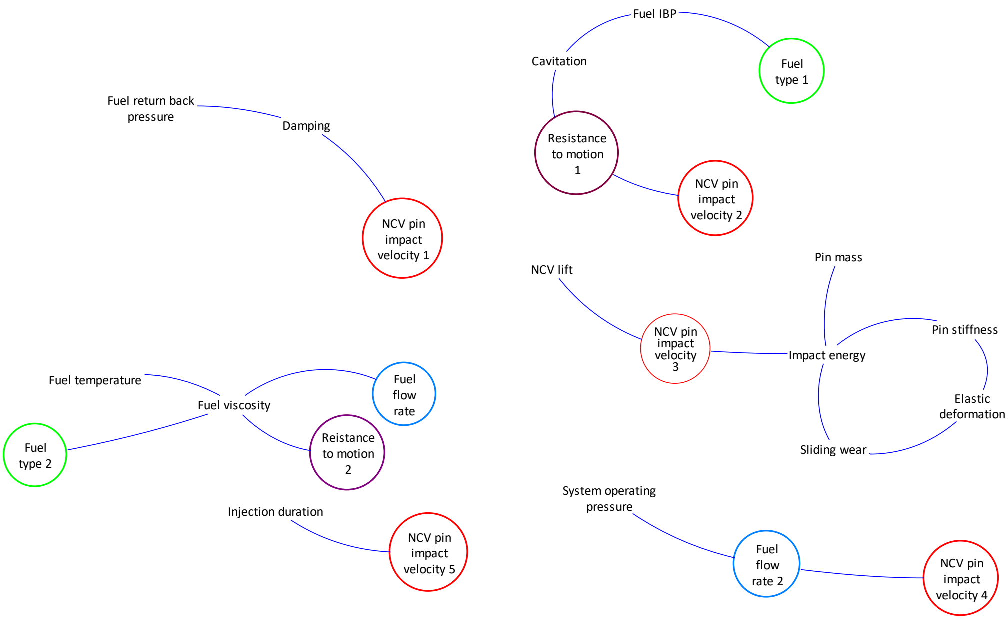


Figure 45: Example of recurring nodes

As discussed in §6.4.3, the quality and quantity of the qualitative justifications provided by the panel members varied significantly, resulting in gaps in logic that translated into disconnected causal fragments. To avoid the influence of bias, the author used only first principles, for instance the relationship between temperature and pressure of a fluid, to connect those fragments rather than their own expert judgement. For example, if the expert panel identified that variable 1 interacts with variable 2, variable 3 then interacts with variable 4, and it is demonstrable through first principles that variable 2 and 3 are proportional to each other, then it is possible to connect the variables such that 1-2-3-4.

In order to translate the links and nodes identified by the study, and the application of first principles into a useful system model, causality was then attributed to each link. Through analysis of the qualitative results of the Delphi Study, and through the first principles that describe the relationship between variables, causality was assigned to the links by the author, and translated into Causal Loop Diagrams using Venesim. This process captured not only the variables that influence NCV seat wear, but also how NCV seat wear manifests as a failure mode in the field

6.5.2 Developing the system model

The resultant system model, in the form of a CLD, was then developed through a number of design iterations, with the dual objectives of capturing the rich detail associated with the study, while retaining a form that provides maximum value when used as a boundary object. A reduced form of the model is also generated such that it can be readily transformed into a parameterised model of the failure mode, representative of the variables determined to be most significant to NCV seat wear. This process was largely completed after the completion of the Delphi Study, with engagement with key project stakeholders and individual members of the expert panel. The results of this process were then communicated back to the entire expert panel.

The first iteration of the system model was generated using the qualitative feedback, in the form of descriptions and comments, provided by the expert panel during the Delphi Study. Figure 46 shows this iteration, with node polarity identified. In this model, and subsequent iterations, the node associated with NCV seat wear itself is highlighted for clarity purposed, alongside the possible end result of the 'Check Engine Light' being illuminated[1]. Furthermore, dependent transition variables not directly identified through the Delphi study, but used in the qualitative feedback provided by the expert panel in their description of the system, are differentiated from those variables identified in the study. This iteration of the model includes a total of 35 variable nodes, including 20 transition variables, and 43 interactions. Furthermore, the model features two reinforcing loops associated with

[1] Figures 46 through 49 are presented out of line with the associated text such as to allow them to be printed in landscape to improve legibility.

the Seat Wear node, describing how existing wear has the potential to lead to an increase in future wear through the increase of effective lift, and an increase in the flow area across the seats.

The second iteration of the system model introduces the complete set of variables and interactions identified through the formal element of the Delphi Study. For clarity, Figure 47 differentiates those newly introduced variables through the use of colouration and style of the associated nodes and links in the model. This iteration re-introduces the 20 variables, with an additional 35 associated interactions included.

The third iteration of the system model is a result of direct stakeholder interaction. Through discussion of both the first and second iterations, it was proposed that an effective way of communicating the model would be to group the variables by where in the project lifecycle they feature, using the Design, Manufacture, and Usage distinction employed in round 1 in the Delphi Study when the variables were originally identified by the expert panel. This was achieved through the colourisation of the text associated with each variable node as shown in Figure 48, with variable groups being largely physically centred in the model also.

The fourth iteration of the model includes the relative significances of each variable and interaction as identified through the Delphi Study, adding another layer of detail. In this iteration, the existing system model is extended to include graphical representations of the significant associated with each variable and interaction as shown in Figure 49. In this iteration, the significance of each variable is identified through the style of node used: variables for which a consensus was reached as being of high significance to NCV seat wear are identified using a circular node, while variables for a which a consensus was reach as being of low or no significance to NCV seat wear are identified using italicised font, in a reduced size. Furthermore, the level of consensus reached for the significance of each interaction is represented through the use of different line thicknesses for the links, with thicker lines indication a higher level of consensus on the significance of an interaction. This iteration of the system model provides a rich level of detail, capturing the full results of the Delphi Study, both with respect to the direct outputs, and the additional detail captured through descriptions of the failure mode provided by the experts. However, the level of detail provided can also act as a barrier in its use as an effective boundary object.

This iteration of the system model provides a rich level of detail, capturing the full results of the Delphi Study, both with respect to the direct outputs, and the additional detail captured through descriptions of the failure mode provided by the experts. However, the level of detail provided can also act as a barrier in its use as an effective boundary object.

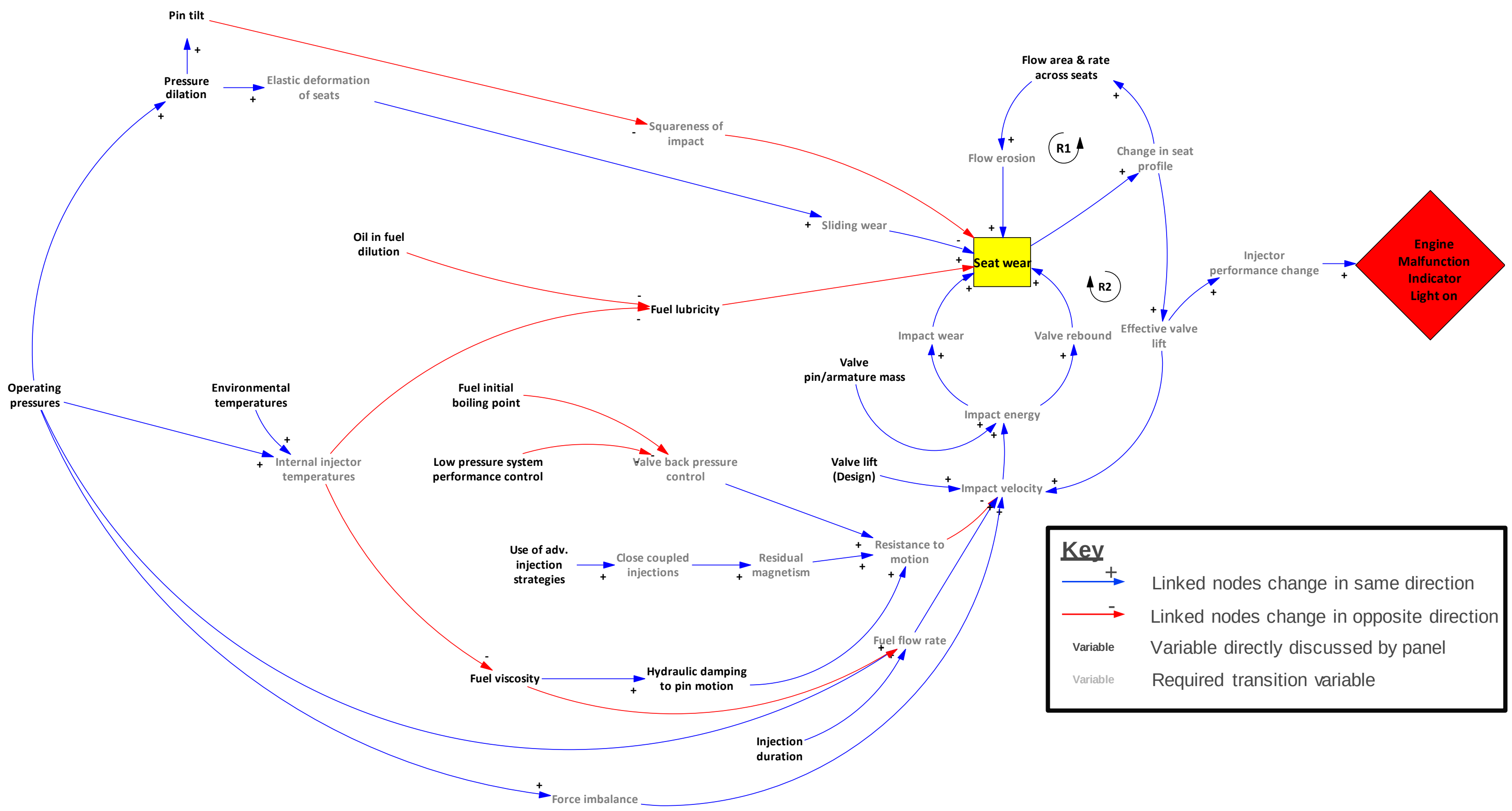


Figure 46: Original system model from qualitative feedback

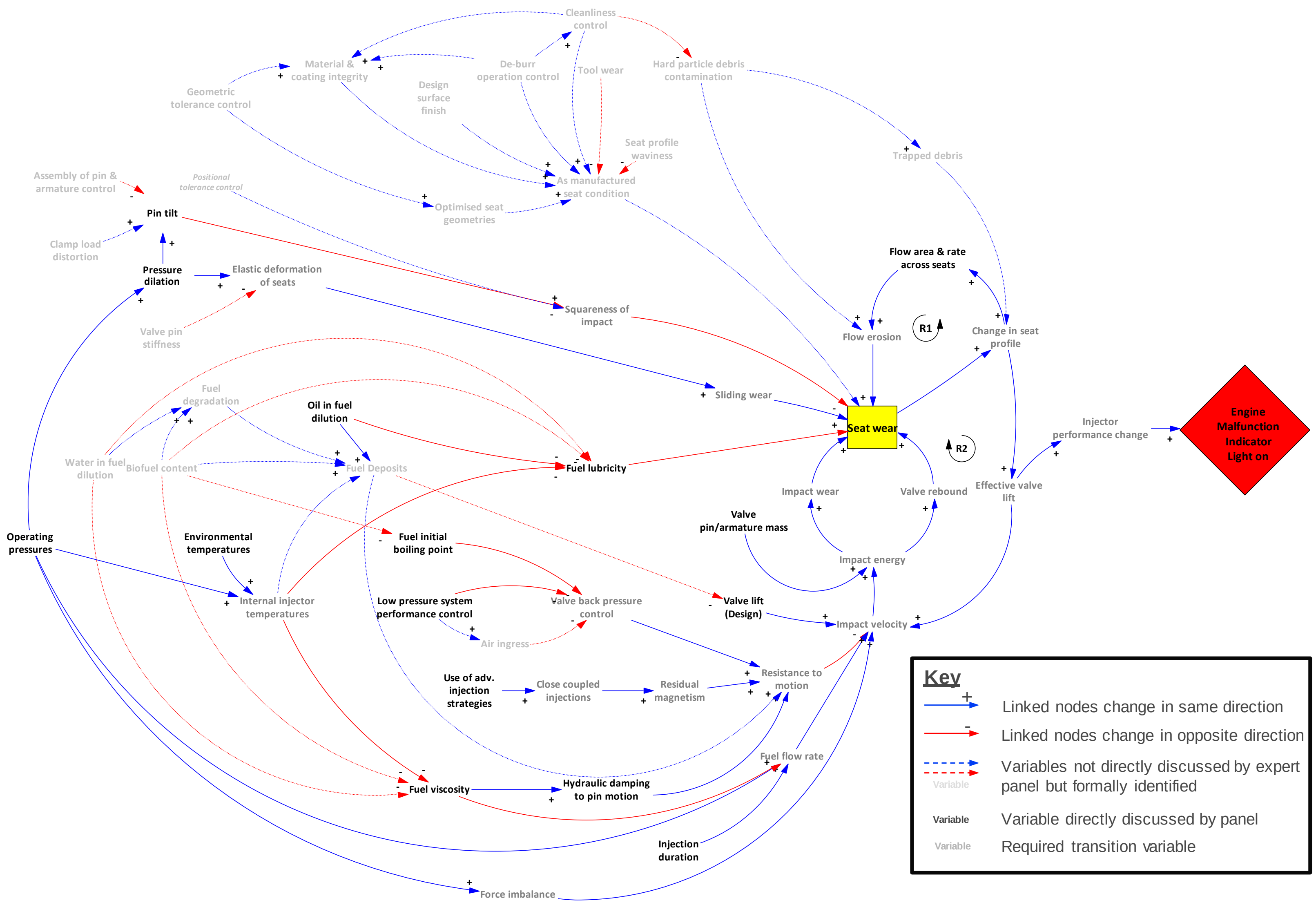


Figure 47: Variables from structured element of the Delphi study re-introduced

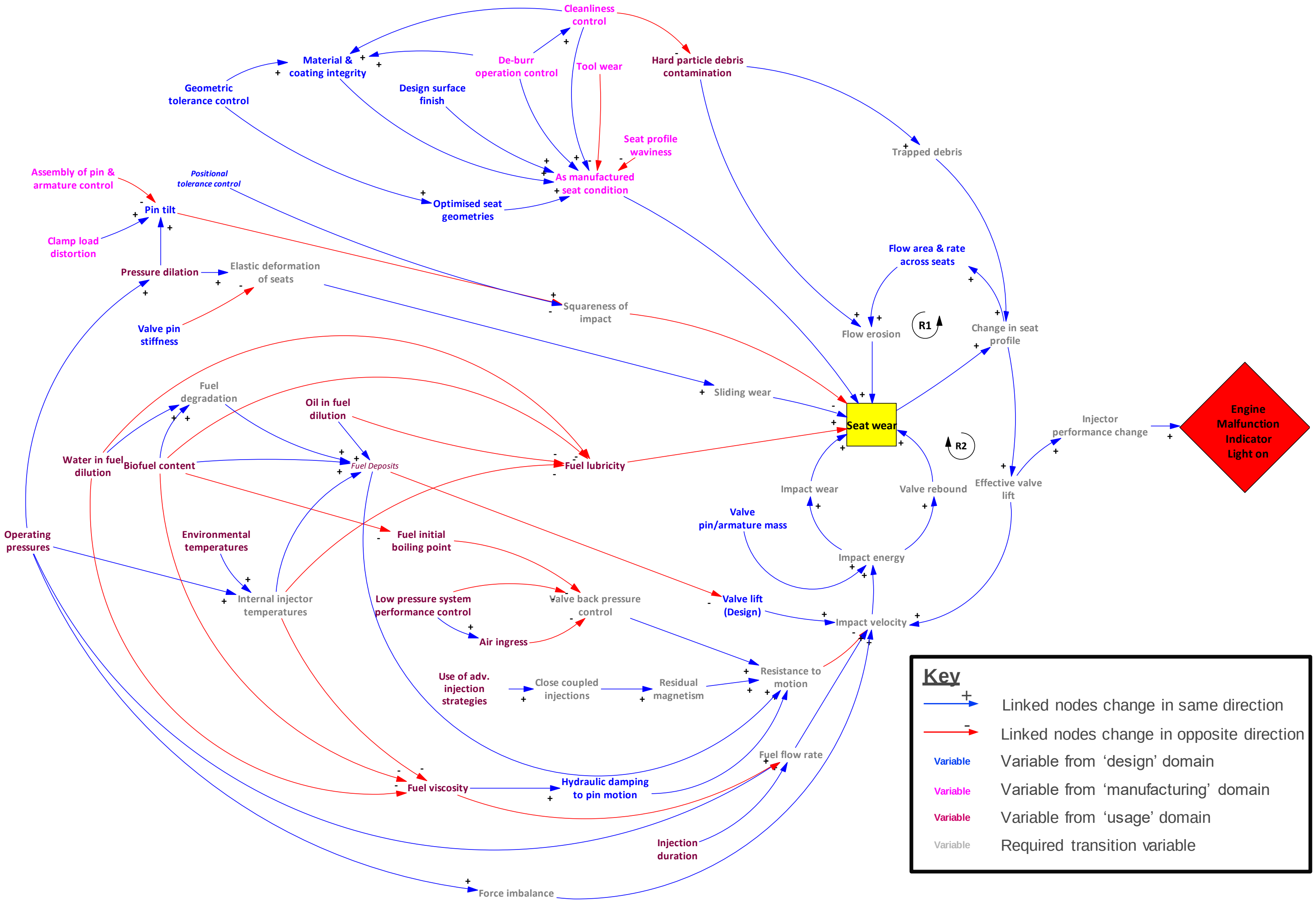


Figure 48: Variables nodes colourised by domain

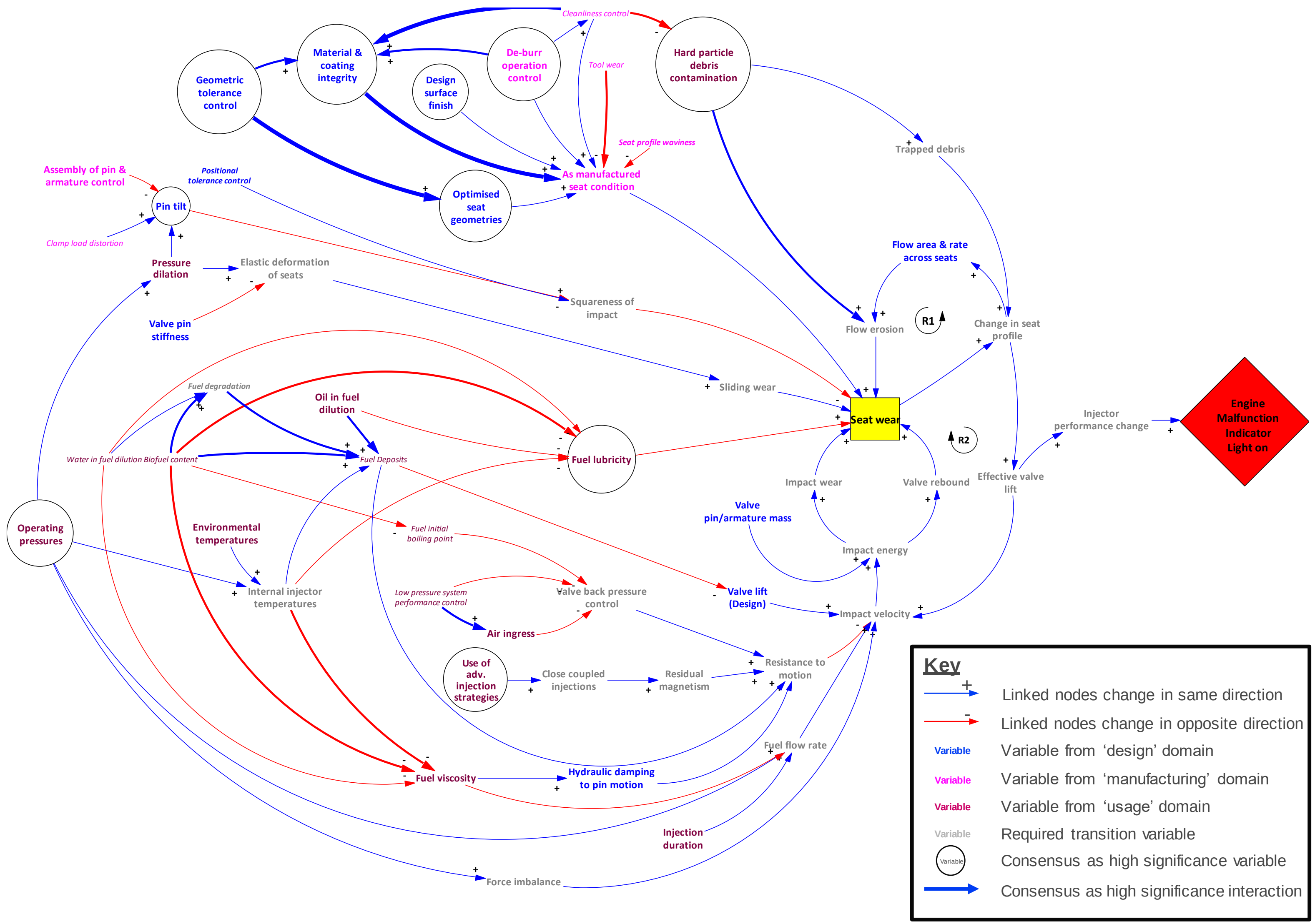


Figure 49: Visual representation of nodes where consensus was met as being of high significance

Figure 50 shows the progression of the model building process, providing an overview that demonstrates how the rich detail in the model increased through iteration.

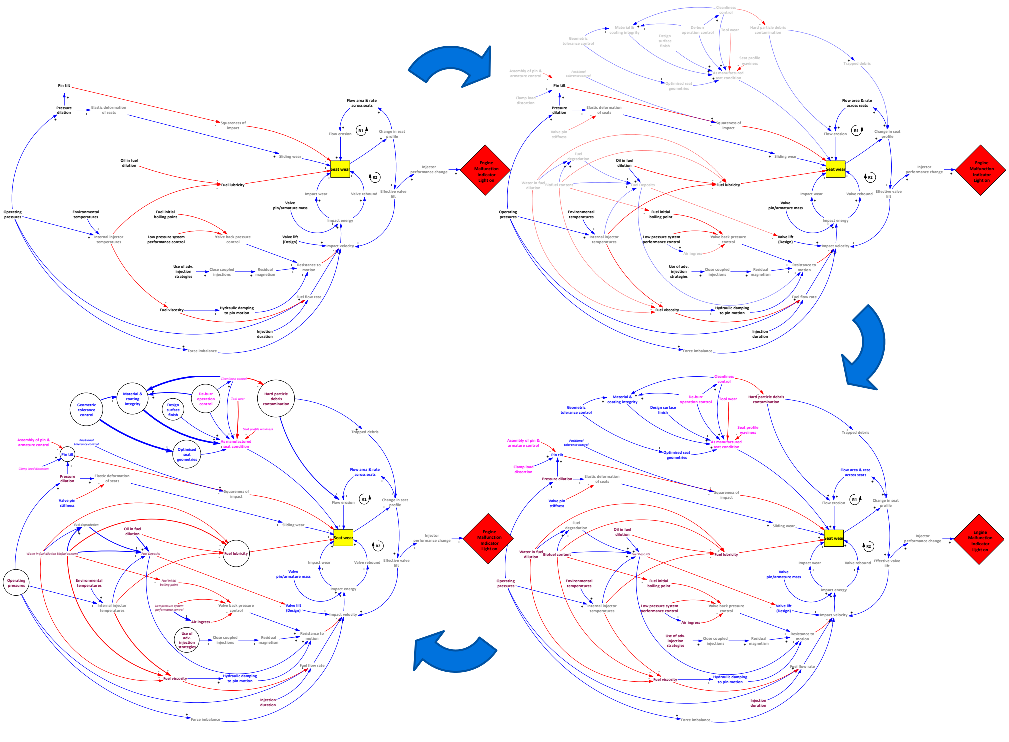


Figure 50: Detail addition within the model building process

In parallel with the objectives of the Delphi Study, the model building process then turned towards reduction of the model, capturing only the most significant variables in order to facilitate the utilisation of the model both as a boundary object, and as an input to the analytical modelling process.

The first iteration of the system model in this reduction phase removed the variables identified through the Delphi Study as being of either of low, or no significance to NCV seat wear. Furthermore, variables where no consensus of their significance was met, were only retained if they had a dependant relationship with variables were consensus was met as being of high significance to NCV seat wear. Transition variables were retained as appropriate in order to best describe the failure mode. Figure 51 shows the result of this iterative reduction in the model. Through this iteration, the model is reduced to 42 variable nodes and 52 interactions.[2]

The final iteration of the model in this reduction phase includes only those variables where consensus was met as being of high significance to NCV seat wear, and the dependent transition variables that describe their relationship with NCV seat wear. The result of this iteration is shown in Figure 52. This version of the system model contains 35 variable nodes and 48 interactions.

[2] Figures 51 and 52 are presented out of line with the associated text such as to allow them to be printed in landscape to improve legibility.

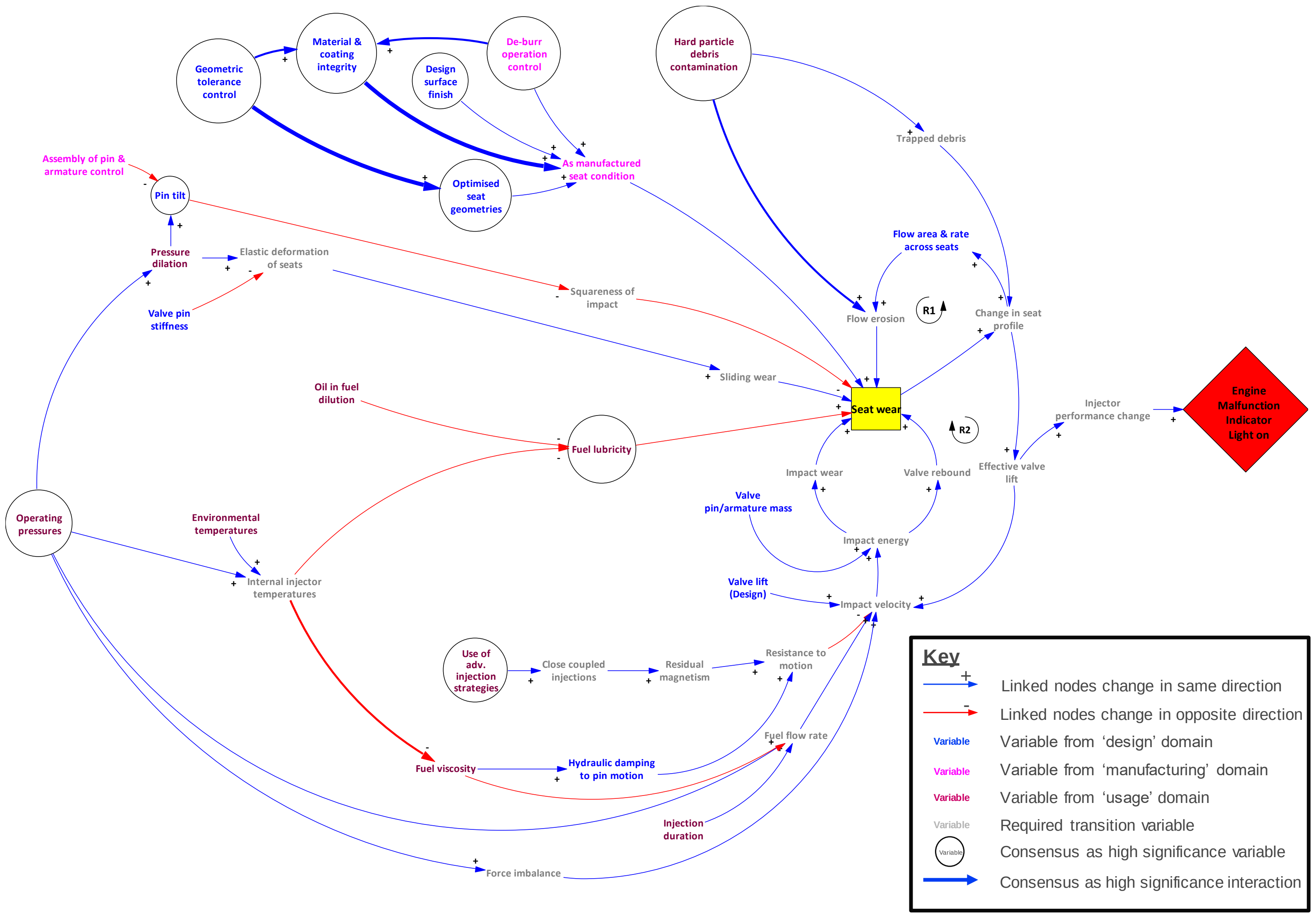


Figure 51: Reduction of nodes where consensus was met as being of low-to-no significance

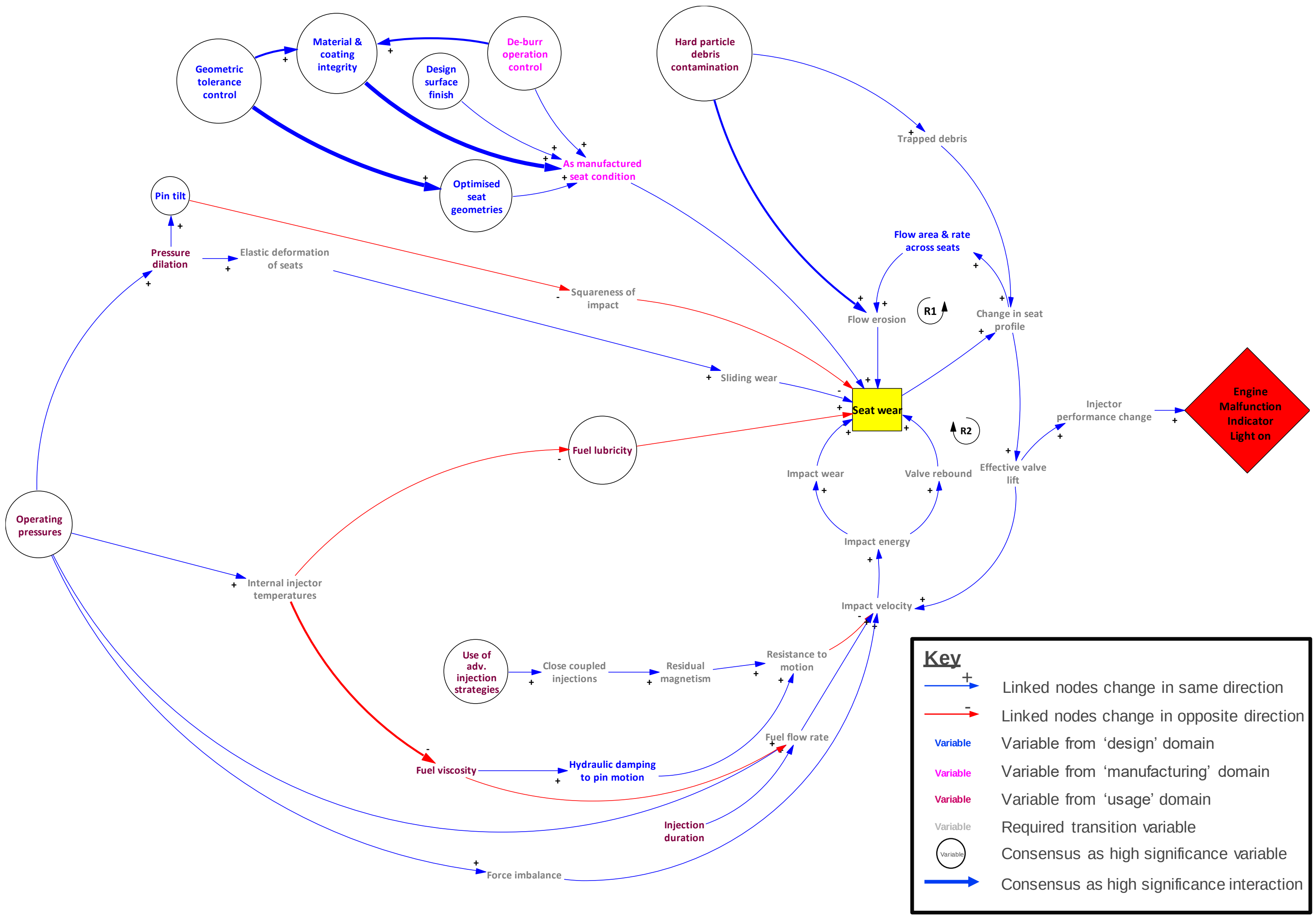


Figure 52: Reduction to only nodes where consensus was met as being of high significance

Figure 53 provides a visual summary of the model building process, comparing the model originally generated from the qualitative feedback of the panel, with the result of the detail addition phase that captured the full result of the Delphi Study, and the result of the reduction phase, with a model intended to describe only the most significant influences on the failure mode.

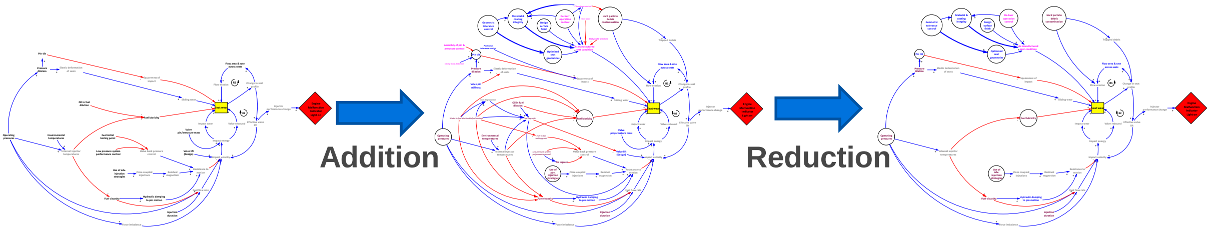


Figure 53: Visual summary of the model building process, from addition through to reduction

While the original model and the final reduced model both contain 35 variable nodes and a broadly similar structure, there are some key detail differences between the two models. A total of 6 variables for which consensus was met during the Delphi Study as being of high significance to NCV seat wear were excluded from the original model. Of these six variables, 4 concerned the ‘Design’ phase of the product lifecycle, 1 concerned ‘Manufacture’, and 1 concerned ‘Usage’ as summarised in Table 7. As can be seen, the original model failed to include any of the variables associated with the manufacturing phase of the associated product lifecycle.

Variable Name	***Phase of product lifecycle***
Geometric tolerance control	*Design*
Material & coating integrity	*Design*
Design surface finish	*Design*
Optimised seat geometries	*Design*
De-burr operation control	*Manufacture*
Hard particle debris contamination	*Usage*

Table 7: High significance variables omitted from original system model

More generally, the original system model is largely comparable to the final model with respect to structure and resolution. The qualitative feedback provided by the expert panel in their description of the failure mode and the variables that influenced it, therefore provided a useful approximation to the system, while presenting an effective boundary object. The complete knowledge associated with the Delphi study added rich detail to the system model, whilst providing additional insights into the perception of certain high significance variables by the expert panel. However, that rich detail has the potential to limit the effectiveness of the model as a boundary object. Further development of the model, leaving only the most significant variables, resulted in a model that was both an effective boundary model, and a purposeful representation of the system behaviour.

6.5.3 Soft system causality as a validation of the research programme

The soft system causality model was constructed using the results of the Delphi study presented in §6.4. The qualitative results were analysed and coded into causal fragments using the method presented in §6.5.1. However, due to the nature of the system being modelled, and the resultant lack of underlying first principles, any gaps in the model resulting from the process had to be filled using the expert judgement of the author.

The model that describes the soft system causality again took the form of a Causal Loop Diagram and went through a series of design iterations to optimise its usefulness as a boundary object. This process was completed outside of the Delphi Study with engagement with the key project stakeholders, with feedback sought from individual panel members as appropriate.

The first iteration of the model that describes the soft system causality is shown in Figure 54. This model was generated using the qualitative responses provided by the panel members as described in §6.4.5. The panel were asked to comment on their perceptions of the deficiencies of the current sociotechnical system, and as such, the feedback provided resulted in a system model that described the system a function of three perceived short fallings. The short fallings identified by the panel were: a lack of a centralised, systematic approach to problem solving at a business level; poor knowledge management system performance; and a perceived lack of focus on scientific first principles and experimental design. These short fallings, how they combine to result in a lack of a thorough understanding of the failure mode, and the resultant ability for the business to make informed decisions, are highlighted in the figure for clarity of communication, using different node styles. The resultant system model includes a compound balancing loop, identified in the figure.

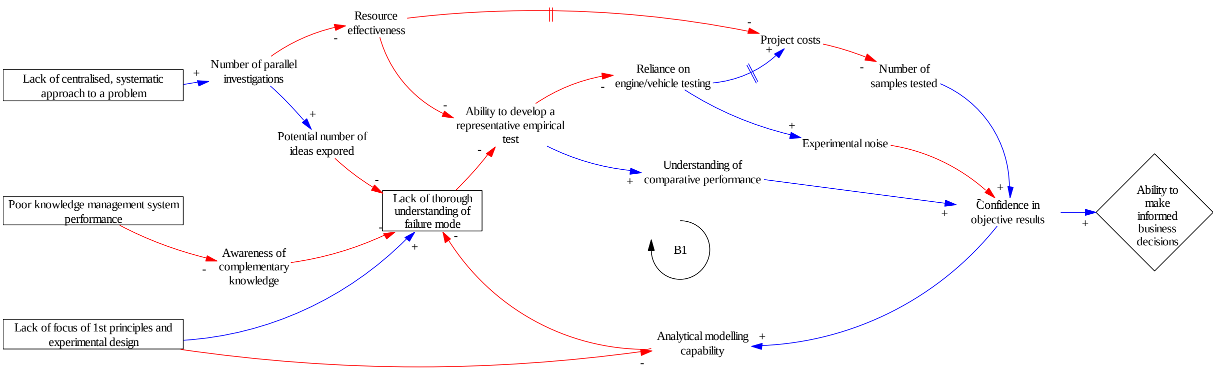


Figure 54: First iteration of the soft system causality model

Interrogation of the resultant system model demonstrates how, without intervention, the three key deficiencies perceived by the expert panel propagate through the sociotechnical system, diluting the available resources such as to increase project costs, ultimately limiting the number of samples available for controlled experimentation. Furthermore, an insufficient understanding of the root cause(s) associated with the failure mode, result in an inability to develop a representative controlled test, which in turn limits the confidence associated with baseline results for any comparative testing, and placing increased emphasis on expensive, and experimentally noisy engine and vehicle testing. A balancing loop in the system then limits any increase in the confidence associated with objective results, resulting in a limited ability for the business to make informed design decisions.

Whilst this iteration of the system model best reflects the responses of the panel to the question asked, it does not adhere to best practices associated with the formation of CLDs. The model contains variables that are 'negative' in meaning, against the recommendation from literature that nodes should feature the positive sense of variables wherever possible to ensure consistency in model formation and interpretation.

The second, and final, iteration of the system model is presented in Figure 55. This version of the system model replaced the previous 'negative' nodes with their positive corollaries. The changes in node polarity results in an associated change in link polarities, and changes the former balancing loop, to a reinforcing loop.

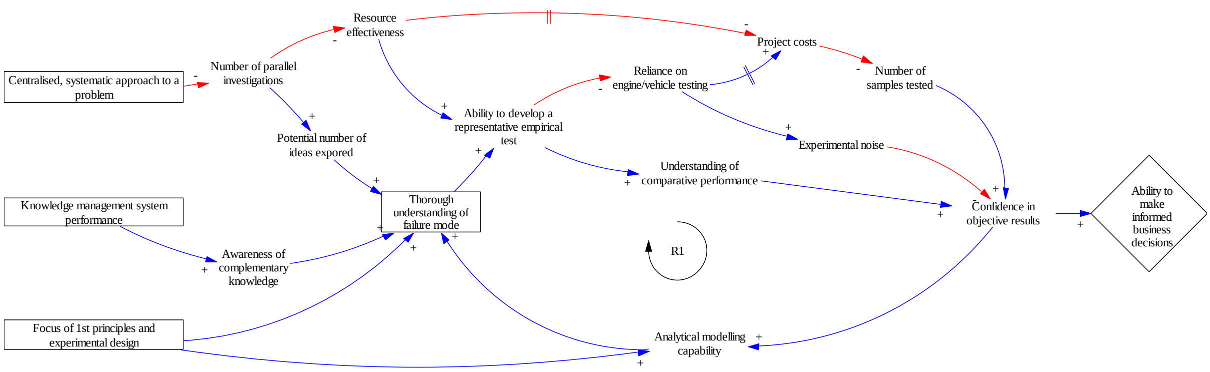


Figure 55: Second iteration of the soft system causality model

In this version of the system model, it can be seen that positive changes in the systematic approach to problem solving, the KM system performance, and the focus on scientific 1st principles, propagate through the system such as to increase the understanding of the root causes(s) of the failure mode, ultimately leading to an increase in the ability of the business to make informed design decisions.

Whilst deficiencies can be perceived in the execution of the first iteration of this system model, feedback from its use with both existing project stakeholders and those new to the problem alike suggest that it is effective as a means for communicating the soft system causality. The second iteration then presents a description of the 'positive' potential end state for the system, allowing viewers to understand how appropriate interventions could improve the business' ability to make informed design decisions. As such, the two iterations of the system model are considered to be complementary.

6.6 Summary

After introducing the Failure Mode that forms the focus of this Thesis, this Chapter has presented an application of the Delphi Method, that Resulted in the exploration, and characterisation of NCV seat wear. A group led definition of the failure mode was agreed by the expert panel, placing emphasis on sliding wear rather than wear associated with valve impact or flow erosion. A total of 35 variables associated with the design, manufacture, and usage of the NCV were identified as influencing seat wear over life. The interactions between those variables were also explored, with the most significant interactions identified. Through iterative exploration of the significance with justifications and evidence shared by the panel, a total of 11 variables were grouped as being at or very close to consensus as highly significant, while 14 variables were identified were informed disagreement was met on their significance. The world views of the panel on the plausibility and desirability of the products being robust to NCV seat wear were also captured, with a consensus that it is desirable, and while it is likely not plausible to be 100% robust, it is plausible to either reduce the wear or the sensitivity of the product such that it no longer remains a significant life wear out mode. Finally, the panel also explored the soft system causality associated with the reasons why NCV seat wear had not been previously fully characterised.

The results of the exploration of the failure mode have been transformed into Systems Dynamic models using Causal Loop Diagram Methodology. The interactions identified by the Delphi study, alongside the qualitative feedback provided by the expert panel, were coded into Causal Fragments through qualitative data analysis. Engineering first principles were then applied to close any gaps in the model as possible, reducing the requirement for any expert judgement bias to be introduced by the facilitator. Finally, the resultant system model went through a number of design iterations upon review with the research stakeholders, improving both its efficacy as a boundary object to be used with different audiences, and its suitability to be transformed into a parameterised, Dynamic Systems model with a lower level of abstraction.

The results of the exploration of the soft system causality that has prevent NCV seat wear from being fully characterised previously have been transformed into Systems Dynamic models using Causal Loop Diagram Methodology. The qualitative feedback provided by the panel, were coded into Causal Fragments through qualitative data analysis. Engineering first principles were then applied to close any gaps in the model as possible, reducing the requirement for any expert judgement bias to be introduced by the facilitator.

The next Chapter will provide an overview of the second element of the FMC process associate with this thesis, using the codified expert judgement from Chapter 6 as inputs into the design of an empirical investigation.

Chapter 7 NCV Seat Wear – Characterisation through Empirical Testing

7.1 Introduction

This Chapter introduces the second element of the case study of the Failure Mode Characterisation method outlined in this thesis and represents the core of this research. In this element, the expert judgement codified in Chapter 6 is used as the inputs for a structured empirical investigation using Experimental Design methodology. The variables identified in Chapter 6 are reviewed for their suitability as design factors in an experimental study, with a focus on usage variables such as to generate a representative accelerated test for future assessment of robust design solutions. Additionally, response factors are identified and developed as appropriate. A pair of iterative experiments are presented, serving to inform the final experimental design and variable selection. A Fractional Factorial design is presented as the selected methodology in this case study, with three design factors, each tested at two levels, resulting in a 2^{3-1} Fractional Factorial design. In each treatment combination, a total of 12 samples are tested for 1000 hours, with performance characterisation performed at 0, 500, and 1000 hours.

The results of the empirical investigation are provided, identifying the usage variable that results in a significant effect in the physical wear to the bottom seat, alongside other main effects and interactions observed in each response variable. The injector performance metrics are then discussed further, identifying which of the existing and new metrics best infer bottom seat wear. Furthermore, a summary of the injector performance change over time is presented through the performance characterisation tests. A centre point is then introduced into the experimental design in order to assess the linearity of the most significant effect, the results of which suggest a non-linear response.

The results of this element of the FMC method represent knowledge of the main effects and interactions observed in NCV bottom seat wear and injector performance associated with usage variable, alongside degradation in injector performance over time. More generally, this element of case study presents the use of structured empirical investigation as a means for transforming the results of the expert elicitation element into a statistically significant, numerical understanding of the failure mode that forms the basis for further analytical and empirical investigation. An overview of this Chapter, and element of the FMC method, is visualised in Figure 56.

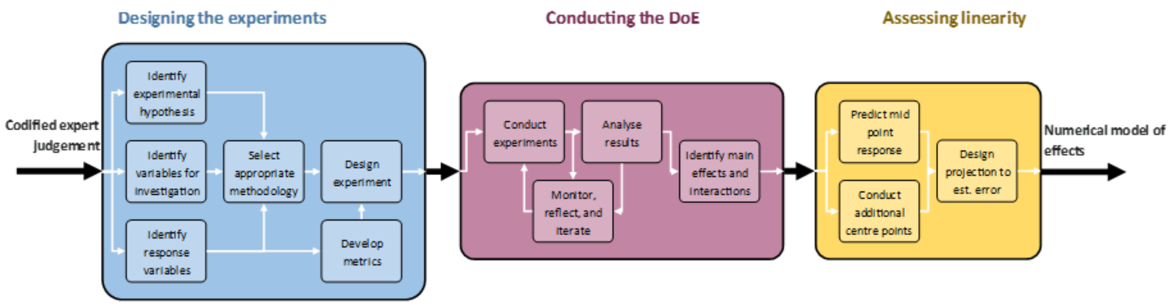


Figure 56: Overview of this Chapter

7.2 Experimental Purpose, Assumptions, and Considerations

The results of the Expert Elicitation study presented in §6.4 were used as the starting point for selection of appropriate variables for this experimental design. The study reached a consensus on 10 variables being of high significance to NCV seat wear, with 5 variables associated with the design of the components, one being associated with the manufacture of the components, and 4 being associated with the usage of the components.

With regards to usage variables, including environmental considerations, a summary of the variables concerning injectors possible to control on test with the facility available is shown in Table 8.

<u>Variable name</u>	<u>Comment (I.e. typical range of variation)</u>
Fuel type	EN590, ISO4113, K34, GRTD, WWLTF, etc.
Fuel inlet temperature	~30degC - ~80degC, dependant on fuel flow
Injection pressure	350bar - 2700bar
Injection strategy	≤ 5inj/cycle; ≤ 900mm3/cycle; injection quantity, timing, & separation
Hard particle debris contamination	Uncontrolled or ISO Class 8-28
Test duration	Typically, 0-4000hrs

Table 8: Variables possible to control on hydraulic test rig

7.2.1 Purpose of the experiment

The fundamental purpose of this experimental study is to determine which of the usage variables identified as part of the expert elicitation study are of significance to the wear of the bottom seat the NCV. This purpose includes the identification of both location and dispersion effects associated with each variable, to determine their influence on both the magnitude of any NCV seat wear, and on the part to part variation in NCV seat wear. As NCV seat wear is believed to be a significant element of performance Drift over Time (Drift), the location and dispersion effects of each variable will be assessed based on both the performance of the injector as measured in a hydraulic test rig, and through subsequent measurement of the NCV seats. Table 10 summarises the primary and secondary purposes of this experiment.

Primary Purpose:	***Identification of the usage variables that are significant to NCV seat wear***
Secondary purposes:	*Generate parametrised regression model of DoT that can compared to the working mental model*
	Identification of suitable metric(s), if any, of injector performance that can be used to infer NCV seat wear
	Provide knowledge that can be used to correlate existing results
	Identify usage variables suitable for design of an accelerated test
	Validate the knowledge of the expert panel and the implementation of the EE method
	Drive capability improvements and promote Experimental Design Methodology within Delphi Technologies

Table 9: Purposes of this experiment

7.2.2 Assumptions

The assumptions associated with this experimental study are as shown in Table 10.

Assumption:	1.	NCV seat wear can be accelerated on a hydraulic test rig using recirculated fuel
	2.	NCV seat wear can be accelerated through variations in usage variables while not being confounded by other failure modes
	3.	The expert elicitation study has identified the usage variables that have the most significance on NCV seat wear
	4.	NCV seat wear can be correlated with injector DoT

Table 10: Experimental Assumptions

7.2.3 Practical considerations

The physical test facility would only be available for a fixed duration as the PV testing associated with the Euro VI family of products phased out, and testing of future products phased in, meaning that effective use of the available resource was an important consideration in the design of this study.

The test facilities available to support this experimental study consisted of one test rig capable of running two full engine equivalent FIE systems in parallel, with the potential for testing 12 samples concurrently. Running two FIE systems in parallel on each test offers a form of partial pseudo-replication. The test rig drives both high pressure pumps through the same gear train, so pump and system speeds will be closely matched through the duration of the test. The fuel supplied to both FIE systems comes from the same tank, meaning fuel properties will be identical for each system, and therefore every injector sample. While the fuel supplied is the same to each system, minor differences in the fuel flow to each system (from differing system efficiencies) and the length and insulation of fuel supply lines to each system will result in some differences in fuel inlet temperature.

For the two parallel systems, for the same pressure demand and injection strategy, differences in the system performance and efficiency will result in slight differences in operating pressures and temperatures. This represents a form of pseudo replication, as some variables that are of potential influence on NCV seat wear will be subject to experimental error as a result of such differences.

Furthermore, for the full test duration, the injector samples will be tested in the same cylinder head pocket. The pocket used for testing will influence the operating pressure for that injector, the axial load applied to the injector, and fluid flows. Pressure wave activity in the high-pressure volume of the rails, pipes, and injectors will result in a variation in the pressures acting on each NCV, but by using the same hardware and rail pressure control strategy as in application, this variation will be within the range typically associated with usage.

Similarly, while the injector clamps and the tightening torque settings used to provide axial load for each injector and pocket combination will be the same, differences in stiffness's of the bolt and the sealing washer used on the bottom of the capnut will result in variation in axial load within each sample population. However, this variation will be within the range typically associated with usage.

By ensuring the injectors are always tested in the same pockets, the back pressures acting on both the injected and leaked fuel will be consistent for each test sample, while in the case of the leakage flow, varying within the ranges typically associated with usage. As the coolant flows through the cylinder head longitudinally, each pocket, and therefore injector sample, will be exposes to a consistent, but varied temperature. This variation will be as typically associated with the effects of engine coolant systems on IIT.

Fuel selection and usage thereof is another area where practicalities must be considered alongside experimental design. An ISO4113 calibration fluid is typically used as a replacement owing to it having similar hydraulic properties to EN590 but has a lower propensity towards combustion and its vapours are less hazardous, and as such, it lends itself to safer unattended running of test rigs. Other fuels come with their own alternative considerations, and those which feature a toluene content to achieve low lubricities have potentially harmful vapours, introducing additional PPE requirements in enclosed spaces such as the test cells.

In order to reduce the costs associated with hydraulic testing of FIE system where no combustion occurs, the test fuel is typically recirculated, with replenishment schedules to account for degradation. In a standard hydraulic test rig, the fuel from injection and leakage is collected, cooled, filtered, and returned to rig tanks. The process of pressurising, and injecting the fuel, subjects the fuels to thermal cycles, exposes the fuels to metals that can act as catalysts for oxidation, allows for cross contamination from lubricant oil or water from condensation, and ultimately has the potential to change the chemical composition of the fuel. This fuel degradation will differ from fuel to fuel and can influence both the hydraulic properties of the fuel, and its propensity for IDID formation, typically accounted for through varying the fuel replenishment schedule. However, when using debris dosing to control hard particle contamination, the closed loop fuel system does not afford the same opportunity for frequent fuel replenishment, so the effects of fuel degradation will be a specific consideration.

In order to best avoid error during the experimental testing programme, close scrutiny and control was to be placed on the test plans and associated procedural documentation and on the experimental running conditions during tests. The tests plans are the main user input to the test rig software, controlling the speed of the drive units, sending commands to the ECUs, and controlling all ancillary systems such as temperature controls and debris dosing systems. Furthermore, the test plans determine what data channels should be recorded, or interrogated from the ECU, and the frequency with which those observations are made. Finally, the test plans are used to place control limits which are used by the test rig software to provide live control over key parameters, allowing the test rigs to run unattended with automatic shutdowns in the event of breaches of those control limits. As such, it can be clearly seen that the role of the test plan is key in the definition, and control of the experimental study.

Another practical consideration in the design of this experimental study concerned the measurement, or inference, of NCV seat wear through the duration of each experiment. Continuous measurement of NCV seat wear would be practically impossible without significantly influencing the system itself. However, several means for continuous and discrete inference of NCV seat wear were considered for this experimental study. The first would utilise observation of the NCV pin motion through a fibre optic camera routed through the injector body. Doing so would allow for continuous inference of NCV seat wear through changes to the effective NCV lift. Whilst this method has previously been proven empirically in short, highly controlled single sample experiments, it had yet to be proven as suitably durable for a multi-cylinder, long hour test environment. Furthermore, the optical observations require significant post-processing to translate into NCV lift, meaning that truly live inference of NCV seat wear would still prove difficult in practice.

An alternative continuous means of inferring NCV seat wear in operation would be the utilisation of radioactive tracers in the coating of both wear surfaces of the NCV bottom seat. Through capturing the injected fuel, and observing the contamination rate of radioactive tracers, wear of the NCV seat could be inferred over time. Doing so would involve considerable additional cost to the experimental programme through the use of bespoke samples, and possible increased fuel requirements through the inability to recirculate fuel in the same manner. Furthermore, the introduction of radioactive tracers to the coatings would represent a disruption to what is a highly controlled system, altering the coating specifications, and potentially impacting coating process performance, coating adhesion, and the performance of the coatings with respect to hardness and wear. As such, the use of radioactive tracers was not considered suitable for this experimental study.

A discrete means of measuring NCV sear wear through the duration of the experimental study would be through the use of physical measurements at prescribed intervals through the testing programme. By removing the injector samples from the test rigs at defined intervals, it would be possible to disassemble the injector, and perform measurements on the NCV. Doing so would require injector performance characterisation both before being stripped for measurement, and after being rebuilt, to quantify any effects of rebuild repeatability on the performance of the injector, introducing additional resource requirements to the experimental study, and increasing the length of time required to complete the programme. Additionally, despite stringent controls, tolerances in assembly of the components would be such that there would be the probability of different axial loads acting on the injector for each build, and that the bottom seat interface between the NCV pin and Piston Guide face would be inconsistent, introducing the possibility of significant experimental error. As such, the decision was made to not use discrete measurement of the NCV seats as part of this experimental study.

Finally, a discrete means of inferring NCV seat wear through the duration of the experimental study is available by characterising the injector performance at intervals through the test programme. By scheduling performance testing to be conducted at defined and consistent intervals through the duration of the experimental study, the performance change over time for each injector, and the sample populations associated with each test can be assessed. While no complete empirical model was available to infer NCV seat wear directly from injector performance characterisation, it had been proven that there was a correlation between the two, most strongly exhibited in the areas of the injectors gain curve most directly influenced by the control valve performance. Furthermore, it is the injector performance change over time that is of concern to the end users, so a further understanding of the effects of different usage variables on performance change over time, and how that relates to NCV seat wear, would be a significant potential output of this experimental study. Whilst this method does not allow for direct measurement of NCV seat wear over time, performance characterisation performance tests at defined intervals, and a developed understanding of the relationship between seat wear and injector performance, is presented as a practical, non-invasive, and relevant means of assessing the rates of both NCV seat wear and injector performance change over time.

7.3 Selection of Design Factors, and Experimental Design method

7.3.1 Possible variables associated with the design & manufacture of the FIS

Of the variables associated with the design and manufacturing of the components, Figure 49 previously demonstrated that they all, with the exception of 'pin tilt', influence the 'as manufactured seat condition' of the components. This clearly indicates significant interactions between the 5 variables, and potential difficulties in producing, or selecting from a large sample size, test samples that accommodate controlled and isolated variation in these variables. The remaining design variable, 'pin tilt' could be explored through the identification of drawings and subsequent manufacture of test samples that demonstrate a variation in pin tilt outside of the current design specification, but the effects on injector performance and alternative wear modes was not fully characterised at the time of this experiment.

For the purposes of this study, the decision was made to consider all variables concerned with the Design and Manufacture of the samples as being frozen, allowing them all to vary within their normal production capabilities. As previously discussed, the focus of this empirical study was on usage variables such as to generate a representative accelerated test. Controlling variation of design and manufacture factors could significantly increase the resource requirements for the test, mandating low volume sample builds of injectors to be completed, meaning high volume, and high capability processes could not be used, potentially increasing undesired variations. Furthermore, direct comparison with existing testing on hydraulic test rigs, and engines and vehicles was desired by all stakeholders.

7.3.2 Possible usage variables where consensus was met of being of high significant to NCV seat wear

Number of injections

The *use of advanced injection strategies*, with multiple injections per combustion event was another variable where consensus was met among the expert panel of its high significance to NCV seat wear, and is highlighted as an element of the system model in Figure 57. In addition to simply increasing the number of valve events for a given test duration, the *use of advanced injection strategies*, with closed couple injections, increases the potential for *residual magnetism* to influence the *force balance* acting on the valve pin assembly.

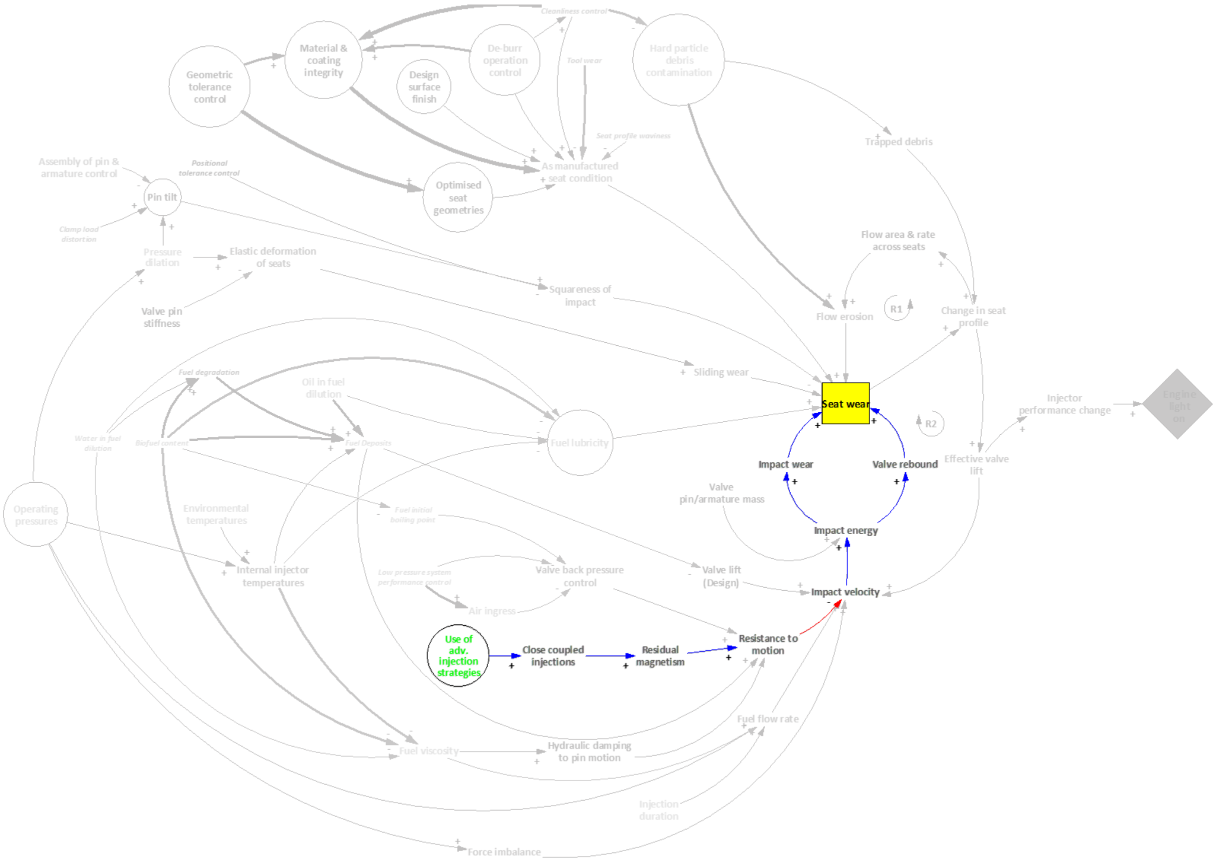


Figure 57: Number of injections as an element of the system model

The *injection strategy* can be controlled consistently and accurately on the test rigs used for this experimental study with respect to timing, duration, and separation of multiple injection events, so the use of *advanced injection strategies*, through controlled the number of injections, could be considered as a possible design factor for this study.

Fuel lubricity

The NCV bottom seat could represent an elasto-hydrodynamically lubricated contact, where the diesel fuel acts as the lubricant. As such, lowering the *lubricity* of the fuel can have an impact on the wear of the NCV bottom seat. Fuel *lubricity* is highlighted as an element of the system model in Figure 58.

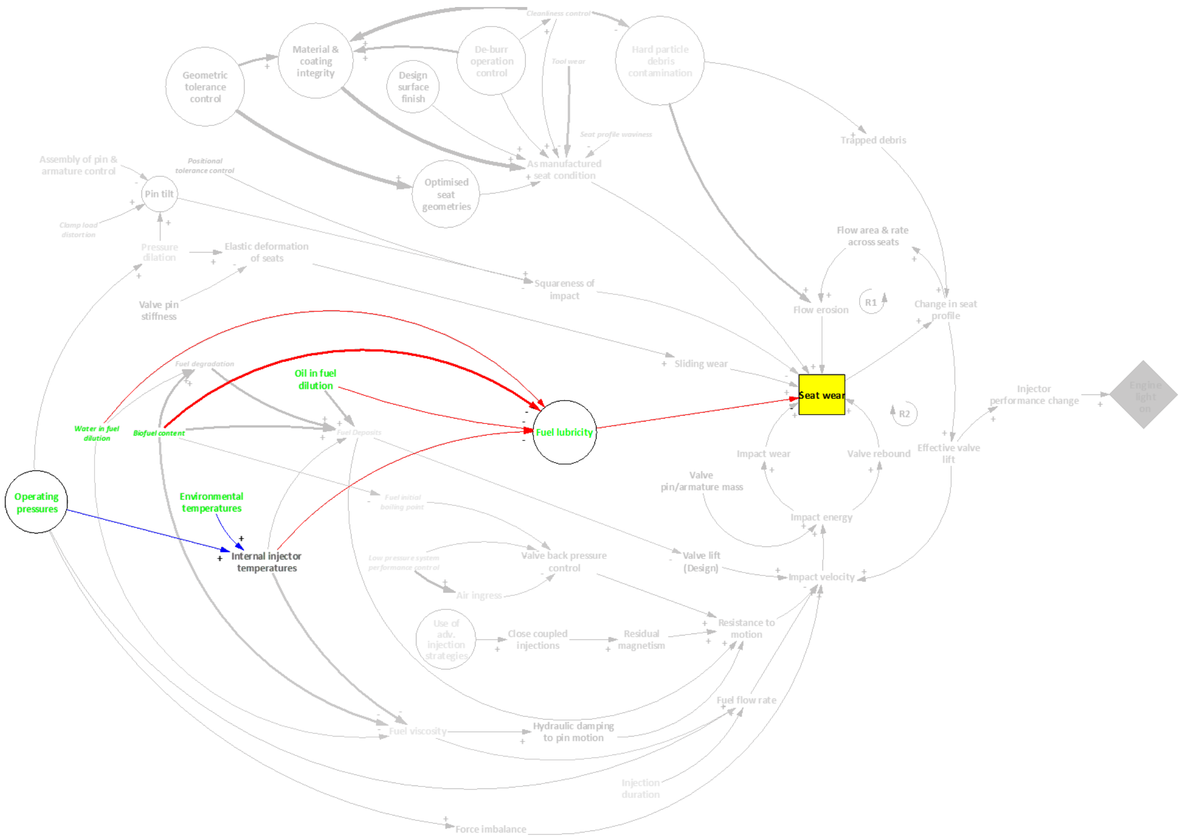


Figure 58: Fuel lubricity as an element of the system model

Fuel properties, such as lubricity and viscosity, can be problematic to vary in isolation. For example, a fuel blend that exhibits low lubricity when compared to EN590, will often also exhibit a relatively low viscosity. Figure 59 provides a visual summary of the comparison of typical fuel specifications with respect to ranges of lubricity and viscosity.

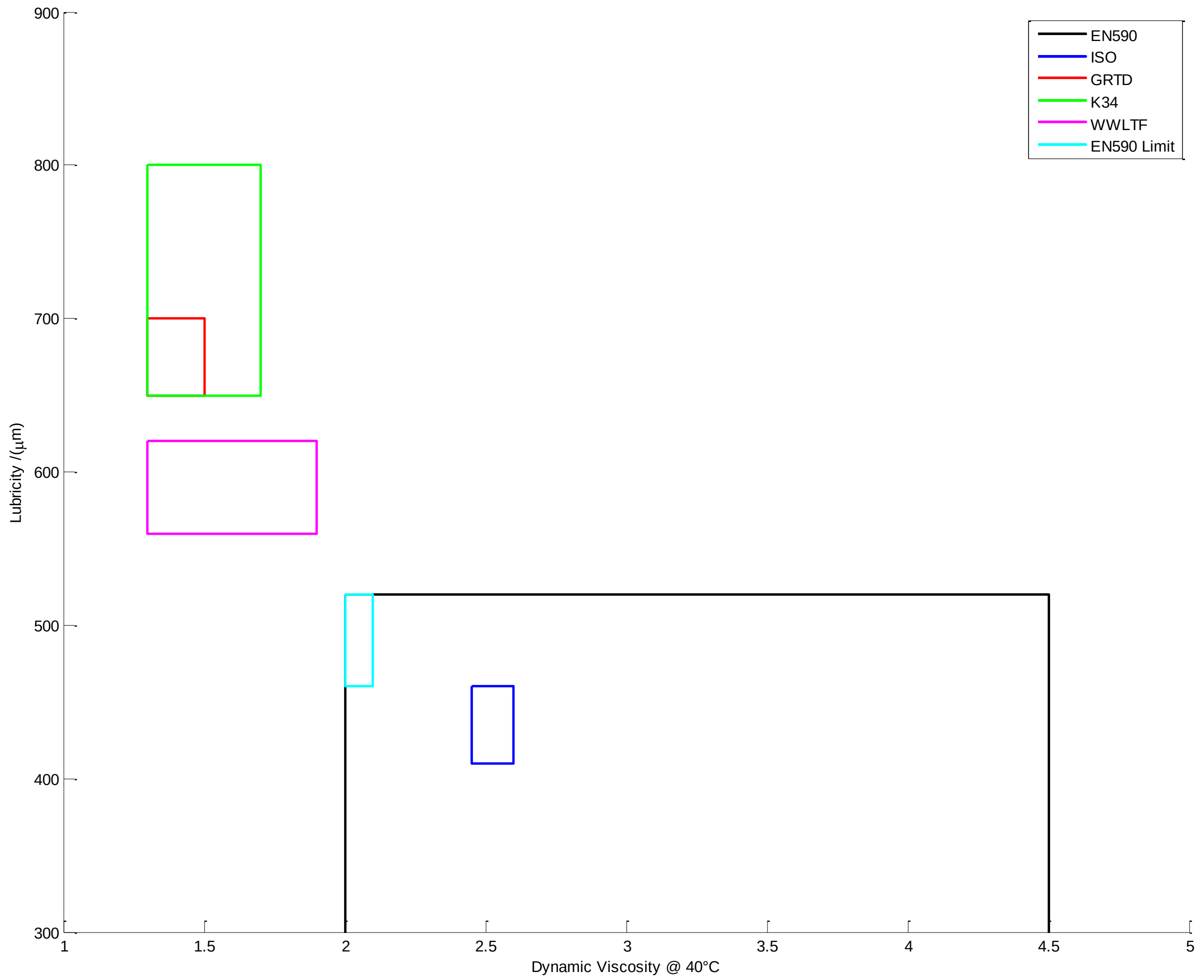


Figure 59: Visual summary of fuel specifications with respect to Lubricity and Viscosity

While it is possible to specify alternative fuel blends that decouple the viscosity and lubricity relationship to a degree, the possible range of that decoupling, alongside the practical considerations in doing so, present significant barriers. For the purpose of this experiment, it was therefore decided to select test fuels from the existing range of fuels used by Delphi Technologies for PV testing. As no consensus was met through the Expert Elicitation study on the significance of fuel viscosity on NCV seat wear, while high significance was agreed for fuel lubricity, it was decided to select fuel based on a range of lubricity values.

Fuels with a lower lubricity have demonstrated a propensity towards accelerated degradation and IDID formation when used on recirculating hydraulic test rigs. Such IDID formation is undesirable with respect to characterising NCV seat wear as it can significantly influence NCV performance, influencing the wear mechanisms this experimental study is intended to characterise. In order to better inform the choice of test fuels, a fuel degradation study was completed as part of the iterative experimentation associated with this study, and will be detailed in §7.3.4.

For the purposes of this empirical investigation, fuel type, can be considered as a possible design factor with an emphasis on fuel *lubricity*. However, it also needs to be noted that fuel *lubricity* has been demonstrated to have a possible variance with fuel temperature, which may represent an interaction with internal injector temperatures.

Hard particle debris contamination

The expert elicitation study reached a consensus that *hard particle debris contamination* is of high significance to NCV seat wear, and is highlighted as an element of the system model in Figure 60. The levels of such contamination are a function of the cleanliness control associated with the manufacture of the components, the cleanliness of the fuel as used in field, and the efficiency and adherence to the required servicing of the fuel filtration system as fitted to the vehicle. The system model generated as part of the expert elicitation study demonstrates how increases in the level of *hard particle debris contamination* could influence both the significance of any *flow erosion* across the valve seats, and the possibility of *trapped debris* resulting in changes to the seat profiles.

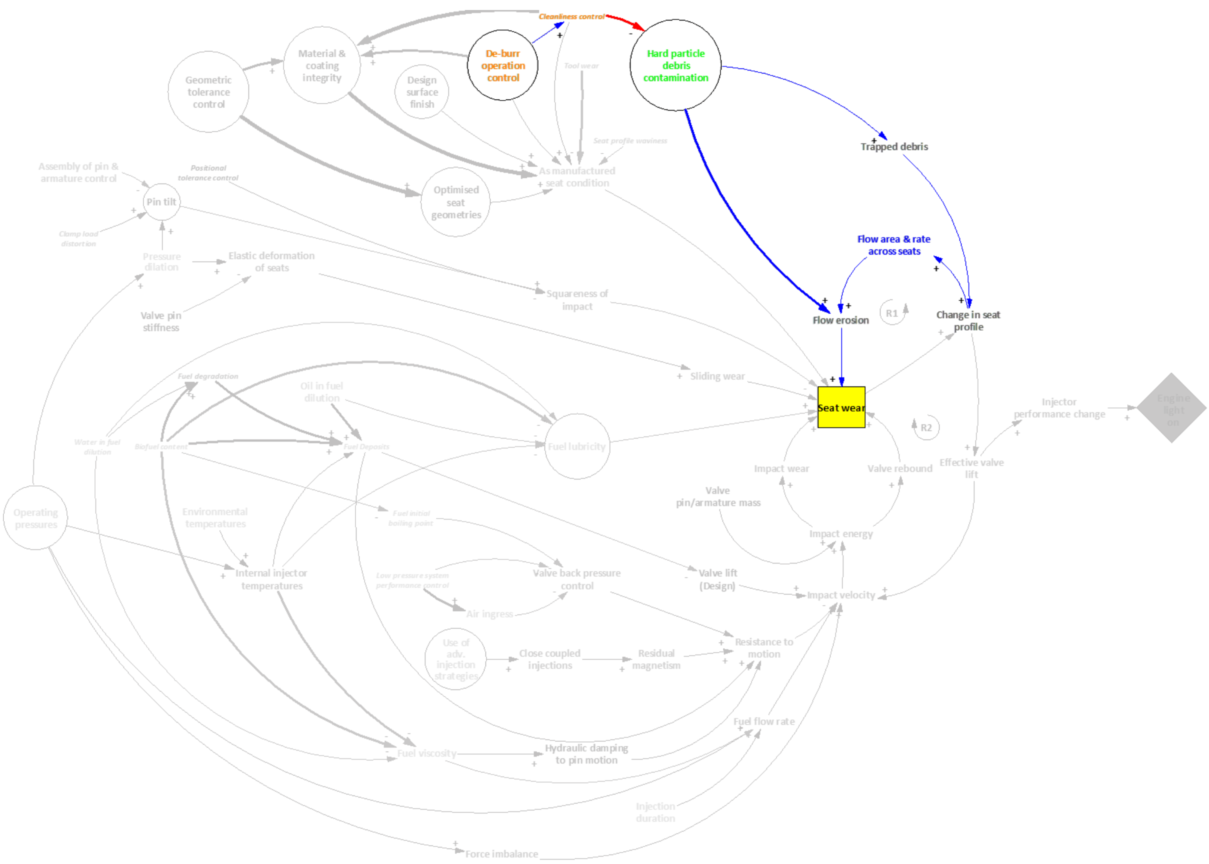


Figure 60: Hard particle debris contamination as an element of the system model

Varying the level of this usage variable, through controlling the concentration of hard particles in the fuel, can potentially impact the validity of the tests with respect to NCV seat wear. Empirical testing has demonstrated that increasing the concentration of hard particles in the fuel results in an acceleration above that associated with total particle count alone, with the potential for significant damage to the valve seat and stem, with dramatically increased leakages. That leakage can influence the performance of the valve with respect to the other variables associated with NCV seat wear, resulting in varying times, and therefore valve actuations to failure, introducing the potential for disparity between tests ran to the same duration. Furthermore, seat damage associated with *flow erosion* or *trapped debris damage*, is not included in the definition of NCV seat wear derived in the Expert Elicitation study. As such, a decision was made to not use *hard particle debris contamination* as a design variable in this experimental study.

However, in being both tightly specified and controlled, and used in a recirculating and therefore highly filtered manner, the fuels typically used for non-fired hydraulic system testing do not provide a representative *hard particle debris level* equivalent to those used in by the customers in application. From the warranted fuel specifications agreed between Delphi Technologies and its customers, and given the associated fuel filtration specifications and efficiencies, it is possible to determine the total number of hard particles the FIE system will be potentially exposed to in application lifetime.

For the purpose of this experimental study, *hard particle debris contamination* could therefore be considered as a possible design factor.

System operating pressure

The *system operating pressure* is a usage variable for which consensus was met in the expert elicitation strategy as it being of high significance to NCV seat wear and is highlighted as an element of the system model in Figure 61. The system model generated through that study demonstrates the influence of the *operating pressure* on NCV wear through both *sliding wear* and *impact wear*. The *system operating pressure* can be controlled accurately and consistently on the available test rigs, using the FIE systems integral pressure sensor, and electronic control valves.

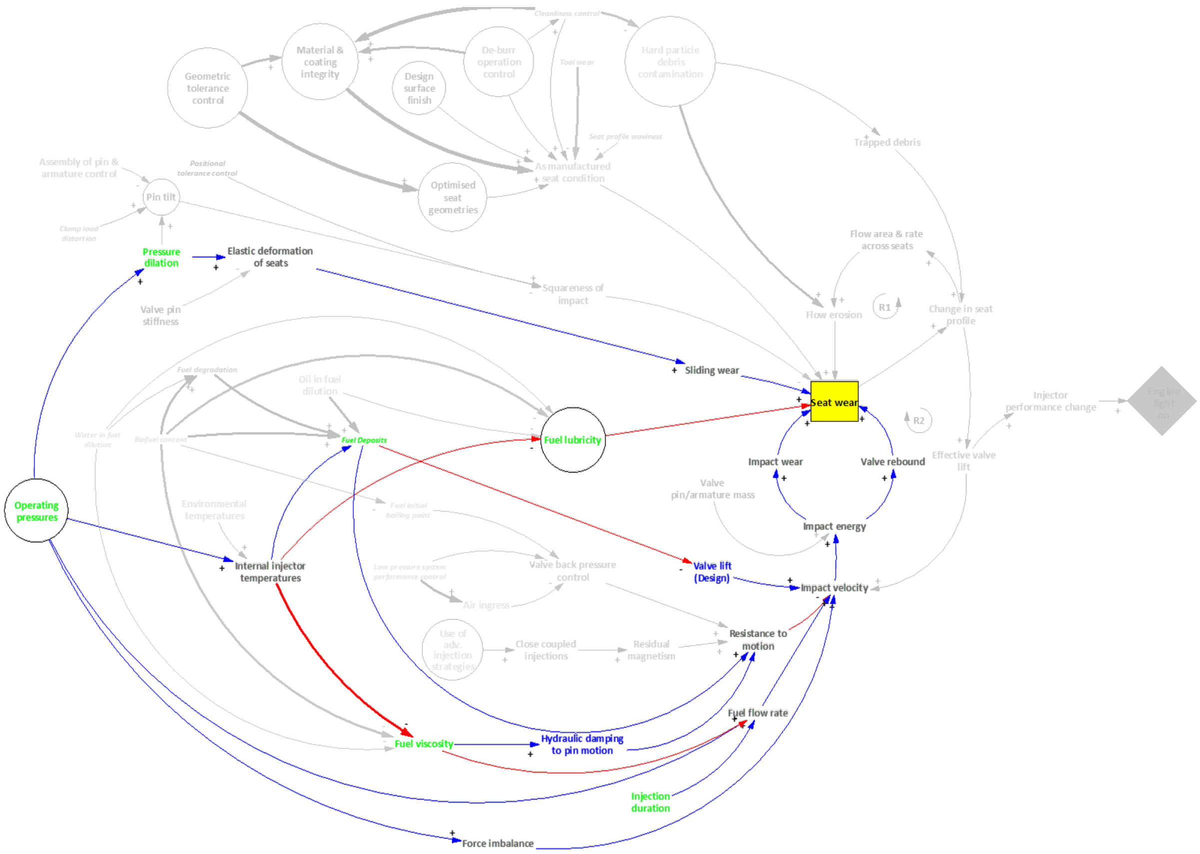


Figure 61: System operating pressure as an element of the system model

Through variations in the pressure around the bottom seat associated with injection events and pressure wave activity, the *system operating pressure* can influence the pressure differential across the valve, and therefore the *pressure dilation* potentially associated with *sliding wear* of the control valve seats. Furthermore, the *system operating pressure* can influence the internal injector temperature, firstly through the laws of thermodynamics, but critically, through increasing the propensity for fuel to leak into relative low flow areas, such as the armature chamber and stem clearance volumes, increasing the propensity for *IDID formation*. As such, increasing the *system operating pressure* can have the effect of reducing the effective *hydraulic damping* through a reduction in the *viscosity of the fuel*, and if it contributes to an Internal Injector Temperature (IIT) associated with *IDID formation*, it can further influence the motion of the NCV pin. Finally, the *system operating pressure* also influences the *force balance* acting on the valve assembly. In addition to the magnetic and mechanical forces acting on the valve assembly through the stator and valve return spring, the pressure balance across the valve is a significant variable in the force acting on the valve, and thus its acceleration and subsequent *bottom seat impact velocity*. As such, *system operating pressure* has a further potential influence on the *energy* associated with bottom seat impact, and thus the *impact wear* associate with NCV valve actuation.

Therefore, *system operating pressure* can be considered as a possible design factor in this experimental study, while recognising it can influence wear through a number of mechanisms.

7.3.3 Usage variables where informed disagreement was met on the significance to NCV bottom seat wear

Oil in fuel dilution

Oil in fuel dilution is a variable associated mainly with the design of the FIE system but can also be influenced in usage. The main mechanism by which this dilution can occur is in oil lubricated fuel pumps through contamination in the plunger bore (or equivalent), typically limited by tolerance control and the feature of wiping edges on the plunger designed to prevent lubricant oil from entering the pumping chambers but can also occur through mishandling of fluids in application. *Oil in fuel dilution* will influence the effective *viscosity* & *lubricity* of the fuel and can increase the propensity for *IDID formation*.

The levels of oil in fuel dilution associated with the design of the system will vary between pumps within a range of variation associated with serial production. Oil in fuel dilution can be monitored during tests by analysing fuel samples for typical elemental markers associated with lubricant oil, but empirical studies on the FIE system used for this experimental study have previously established the typical level and variation associated with production pumps, with both low total levels of dilution and low variation between systems.

As such, and to best ensure the representativeness of this experimental study, this variable is not suitable to be varied as a design factor, instead it will be considered to be fixed within normal production variation.

NCV pin dynamic motion

The *impact velocity* (classified as *NCV pin dynamic motion* by the expert panel) of the NCV valve pin is dependent on a number of design and usage variables. The first design variable to affect velocity is the *valve lift*, where designs that feature higher lifts result in increased *impact velocity*. The designed valve lift, and associated manufacturing tolerance is then potentially influenced by *IDID formation*, where high levels of IDID can result in a reduction in effective valve lift. However, with designed valve lift considered as fixed for the purpose of this experiment, and fuels selected to reduce the possibility of IDID formation, valve lift will not be considered as a design variable to control NCV pin velocity.

The *impact velocity* of the NCV pin is also influenced by the resistance to motion, a transitionary variable itself influenced by the *back pressur*e acting on the drain orifice, any *residual magnetism* associated with the stator, and the *hydraulic damping* associated with armature displacement. The *valve back pressure control* is influenced by the *low-pressure system performance,* any *air ingress*, and the *initial boiling point* of the fuel. The expert elicitation study concluded that *low pressure system performance* is considered to be of low significance to NCV seat wear so will not be considered as a further part of this experimental study. *Air ingress* will be discussed as standalone variable later in this section. The *initial boiling point* of the fuel was determined through the expert elicitation study to be of low significance to NCV seat wear but will be varied in this experimental study through the use of different fuels.

Any *residual magnetism* acting on the valve pin assembly is a product of *closed coupled injections* associated with the use of *advanced injection strategies*. The small pilot injections associated with such strategies occur in the ballistic region of the gain curve, where the end of injection can occur while the NCV drive waveform is still at boost level rather than the low hold level associated with longer actuations, influencing electromagnetic flux decay. *Residual magnetism* will also be influenced by the design of stator drive waveform itself, and the design of the armature, but for the purpose of this experimental study, both will be consider as controlled to within normal production and application variance. For this experimental study, the influence of *residual magnetism* on NCV seat wear will be assessed through varying the injection strategy employed.

The third variable that influences the *impact velocity* of the NCV pin is the *duration of injections* employed, as the *injection duration* will also influence any residual armature forces, hydraulic pressure forces, and mechanical loads acting on the valve, influencing bottom seat impact velocity.

The final variable influencing *impact velocity* is the *force balance* acting on the valve. The *force balance* acting on the valve is a function of the *operating pressure* of the system and the preload and stiffness of the return spring. The return spring preload & stiffness are considered as fixed design variables for the purpose of this experimental study, while the *operating pressure* of the system will be used a design variable.

Due to the multiple influencing variables, NCV pin velocity is not suitable to use as an independent design variable in this experimental investigation, but instead will vary dependently on other variables. Further empirical and analytical investigations could then be designed to determine the effect, if any, of this variable on NCV seat wear.

Fuel viscosity

The system model that resulted from the expert elicitation study provided additional insight into the identification of IIT as an influence to NCV seat wear. As IIT varies, the *viscosity* of the fuel in the armature chamber will change, influencing the resistance to motion associated with hydraulic damping, resulting in a change associated with the kinetic energy of the pin. As the *viscosity* of the fuel decreases as temperature increases, an increase in IIT should then result in an increase in the energy associated with pin impact through a reduction in hydraulic damping.

However, while the *viscosity* of fuel specifications can vary significantly (as shown in Figure x), those specifications reference a *viscosity* as measured at 40ºC as per the ISO test specification. IIT has been demonstrated to vary between ~120ºC for a typical 1800bar application with lower fuel inlet temperatures, and ~200ºC for a typical 2500bar application with higher fuel inlet temperatures. As the relationship between the temperature and viscosity of a fuel is a logarithmic one, fuels that may show significant variation by *viscosity* specification at 40ºC, do not necessarily exhibit significant variance in their kinematic viscosities at the range of temperatures typical of IIT (Lacey et al 2001).

This relationship may suggest that the kinematic *viscosity* of the fuel at representative operating pressures and associated IITs, may not result in significant changes to the hydraulic damping associated with valve actuation. This suggestion is supported by one of the Delphi Study panel members, who while ranking fuel *viscosity* as being of low significance to NCV seat wear said:

> *"In terms of seat wear, main effect of fuel viscosity would be resistance to armature motion, rather than any effects due to seat flow. 99% of NCV operating time is within a small viscosity range."*

When combined with the effects of fuel degradation over time, which as will be demonstrated in §7.3.4 results in fuel *viscosity* decreasing with time at varying rates for different fuel compositions, this reduced range of effective *viscosity* means that viscosity is likely not a suitable design factor.

Environmental temperatures

The *environmental temperature* is a usage variable associated with both the ambient temperature associated with usage, and the installation of the powertrain within a vehicle. The *environmental temperature* will influence the inlet temperature of the fuel to the FIE system, and thus the *internal injector temperatures* associated with application. Typically, on-highway haulage applications are associated with lower *environmental temperatures* than both bus applications, where the powertrain and associated cooling system installation can often be compromised for packaging, and off-highway applications, where both the environmental conditions and powertrain installation can be compromised with respect to *environmental temperatures*. A differential in *IIT* between applications can influence both the effective *fuel viscosity*, potentially the fuel *lubricity*, and the propensity for *IDID formation*, and can therefore influence the *motion of the NCV pin* assembly and the *energy* associated with bottom seat impact. The increased fuel inlet temperatures associated with raised *environmental* *temperatures* can be replicated on test rigs through control of the ambient temperature, the cooling of the return fuel, and through in-line heat exchanges on the fuel supply. FIE system performance at significantly reduced temperatures is a separate concern, and unrelated to NCV seat wear.

Section 7.3.4 will further discuss the *environmental temperature* variable with reference to *IIT* and this experimental study. For the purpose of this experimental study, *environmental temperature* is not considered as a suitable design factor for its potential to influence IDID formation.

Injection durations

Injection duration is a usage variable associated with advanced injection strategies and the duty cycle of the vehicle. As previously described, the injection duration is one of a number of variables that influences the NCV pin velocity, and as such, is not suitable to use a fully independent variable in this experimental study

Air ingress

Air ingress is a usage variable associated with the installation of the FIE system and any associated interfaces with low pressure systems on the engine. *Air ingress* in the low-pressure fuel circuit can influence the *control of the back pressure acting on the drain orifice of the NCV*, with the potential to impact the *resistance to motion* of the valve, and thus the *impact velocity* of the NCV pin assembly. *Air ingress* in the high-pressure fuel circuit could significantly influence valve dynamics and associated *impact velocities*.

Opinion amongst the expert panel varied as to whether a significant deviation in *air ingress*, and *valve back pressure control*, is significant to NCV seat wear, or whether or not it should be a concern to Delphi Technologies as the agreed specifications of associated systems and interfaces with customers dictate it should not occur in usage, nor is it Delphi Technologies' liability. Furthermore, empirical evidence obtained by running a partial vacuum on the NCV backleak drilling resulted in no acceleration of NCV seat wear. As such, the decision was made to not artificially introduce variation in the *air ingress* into the FIE system, or the *control of the valve back pressure* for this experimental study.

Pressure dilation

Pressure dilation resulting in *elastic deformation of the valve seats*, and subsequent *sliding wear* of the NCV seat, is a dependent usage variable where no consensus was met in the expert elicitation study. As previously discussed in §7.3.2, *pressure dilation* is a function of system operating pressure, itself a usage variable where consensus was met as it being of high significance to NCV seat wear. The fact that the expert panel failed to reach a consensus on the significance of pressure dilation alone to NCV seat wear indicates that some consider this element of sliding wear to be of less significance to the sliding wear associated from pin impact, and from impact wear itself. *Pressure dilation* is not suitable to be an independent design factor in this experimental study.

7.3.4 Iterative experimentation

This section provides a summary of two experiments the first of which acted as a pilot for the final experimental design, while the second informed variable selection. Both demonstrate the role of iteration in experimentation.

$2^{(5-1)}$ Half fractional Factorial with Centre Points

An initial iteration of the Design of Experiments study was designed to consider a total of 5 usage variables. While this iteration of the experimental study was not completed, it did provide new knowledge that informed subsequent iterations of the experimental design.

This iteration was conceived in a period of relatively high resource availability, with 4 double headed rigs available for testing. This availability facilitated the inclusion of a wide range of design variables to enable significant exploration of the response associated with each variable. A half fractional factorial design, $2^{(5-1)}$ design was used, as summarised in Table 12, to allow assessment of interactions to be included without confounding at the two-factor level, while centre points were employed to assess the linearity of each variable.

Variable Name	Low Level	High Level	Variable type
Hard particle contamination	Class 0	Class 10	Discrete
Fuel type (lubricity)	ISO4113	GRTD	Discrete
System operating pressure (bar)	1500	2500	Continuous
Fuel inlet temperature (°C)	40	80	Continuous
Number of Injections	2	6	Continuous

Table 11: Design variable of the 2^{5-1} design

The resulting experimental design, including the 4 pseudo centre points required to accommodate the two discrete variables, resulted in a total of 20 test combinations, as shown in Table 13. As the tests would be completed on double header test rigs, and with a view to reducing the total resource and test time required to complete the study, tests were to be ran in parallel wherever possible. It is possible to use different combinations of system operating pressure and injection strategies on each FIE system on a double header rig as long as the fuel related properties remain equal. As such, tests which used the same combination of fuel type, hard particle contamination, and fuel inlet temperature could be ran in parallel. By doing so, the resources required to complete the experimental study would be lower, but there would be a number of considerations with respect to the statistical rigour of the programme. Firstly, the number of samples available for each combination would be limited to 6 injectors. Secondly, by grouping tests by fuel conditions, the randomisation of the design would be reduced significantly. And finally, running tests on rigs in parallel may introduce additional covariant that would need to be assessed in the analysis of the experiment.

Test	Bank	Debris	Fuel	Temp.	Inj.	RP
I	A	0	GRTD	80	3	1500
	B	0	GRTD	80	1	2500
II	A	0	ISO	80	1	1500
	B	0	ISO	80	3	2500
III	A	0	ISO	40	3	1500
	B	0	ISO	40	1	2500
IV	A	0	GRTD	40	1	1500
	B	0	GRTD	40	3	2500
v	A	0	GRTD	60	2	2000
vi	A	0	ISO	60	2	2000
XI	A	10	ISO	80	3	1500
	B	10	ISO	80	1	2500
XII	A	10	GRTD	80	1	1500
	B	10	GRTD	80	3	2500
XIII	A	10	GRTD	40	3	1500
	B	10	GRTD	40	1	2500
XIV	A	10	ISO	40	1	1500
	B	10	ISO	40	3	2500
xv	A	10	GRTD	60	2	2000
xvi	A	10	ISO	60	2	2000

Table 12: Treatment combinations and test groupings in the 2^{5-1} design

Initial testing began following a partially randomised order, before the test samples were removed for performance characterisation and external inspection. Tests with a combination of GRTD fuel, high fuel inlet temperatures, and 2500bar system operating pressure, resulted in externally visible fuel deposits. Furthermore, analysis of the results of this performance characterisation tests exhibited changes typically symptomatic of IDID formation, the presence of which was confirmed through inspection of the test samples. The rate at which the deposits formed was significantly accelerated compared to seat wear and affected the representativeness of the programme. It was therefore decided to investigate IDID formation as a function of the test variables and fuel types such that the DoE programme could be redesigned and restarted before further resource was expended on what would likely be a redundant study with respect to characterising NCV seat wear. The associated IDID experimental investigation is outside of the scope of this thesis but generated new knowledge and test methods for that failure mode, demonstrating the interaction between rail pressure, fuel inlet temperature, and fuel type.

While this initial iteration of the DoE programme was not completed, the process of designing, commissioning, and executing the experimental design in informing subsequent empirical studies, including those directly relating to characterising NCV seat wear and otherwise.

Fuel degradation

In order to better inform the choice of test fuels for further empirical investigations, resource was made available to support a fuel degradation study. In order to allow for testing with several different fuels, and using different replenishment intervals, the fuel sampling interval varied with test duration. Furthermore, while each fuel sample was used for lubricity measurement using internal facilities, the interval between measurements completed by the external contractor was generally increased to limit the cost associated with the study. As such, the test design used for each 300-hour experiment was as summarised in Table 14.

Test duration range/ hours	**Interval between fuel samples /hours** *(External measurements /hours)*
0-48	4 *(12)*
48-192	12 *(24)*
192-300	48 *(48)*

Table 13: Sampling frequency for fuel degradation experiment

For each fuel, where possible, two separate studies would be completed, one using fuel with no replenishment through the duration of the test, and one using the typical fuel replenishment schedule associated with reliability demonstration tests which represented a 20% by volume replenishment every 24hrs. For each test, the relationship between fuel degradation and 6 metrics would be observed, as detailed in Table 15.

Metric	**Comment**
Dynamic Viscosity (cts @ 40degC)	Hydro-mechanical properties
Lubricity (WSD)	
Total Acid Number (mg KOH/g)	Indicators of chemical degradation or contamination
Water Content (mg/kg)	
FTIR absorption (max peak)	
Solids (mg/kg)	

Table 14: Metrics for fuel degradation experiment

For clarity of communication in this thesis, the effects of fuel degradation on the two hydro-mechanical properties will be presented alongside Total Acid Number as a representative indicator of any chemical degradation in the fuel.

With respect to the ISO4113 fuel, the study concluded that the effects of fuel degradation through recirculation were minimal, and that the existing fuel replenishment intervals could be extended significantly while retaining the hydro-mechanical properties of EN590 fuel. Figure 62 shows the observed response of the representative hydro-mechanical and chemical properties of the fuel through the duration of the degradation study, for both use of the standard fuel replenishment schedule, and for the case of non-replenished fuel.

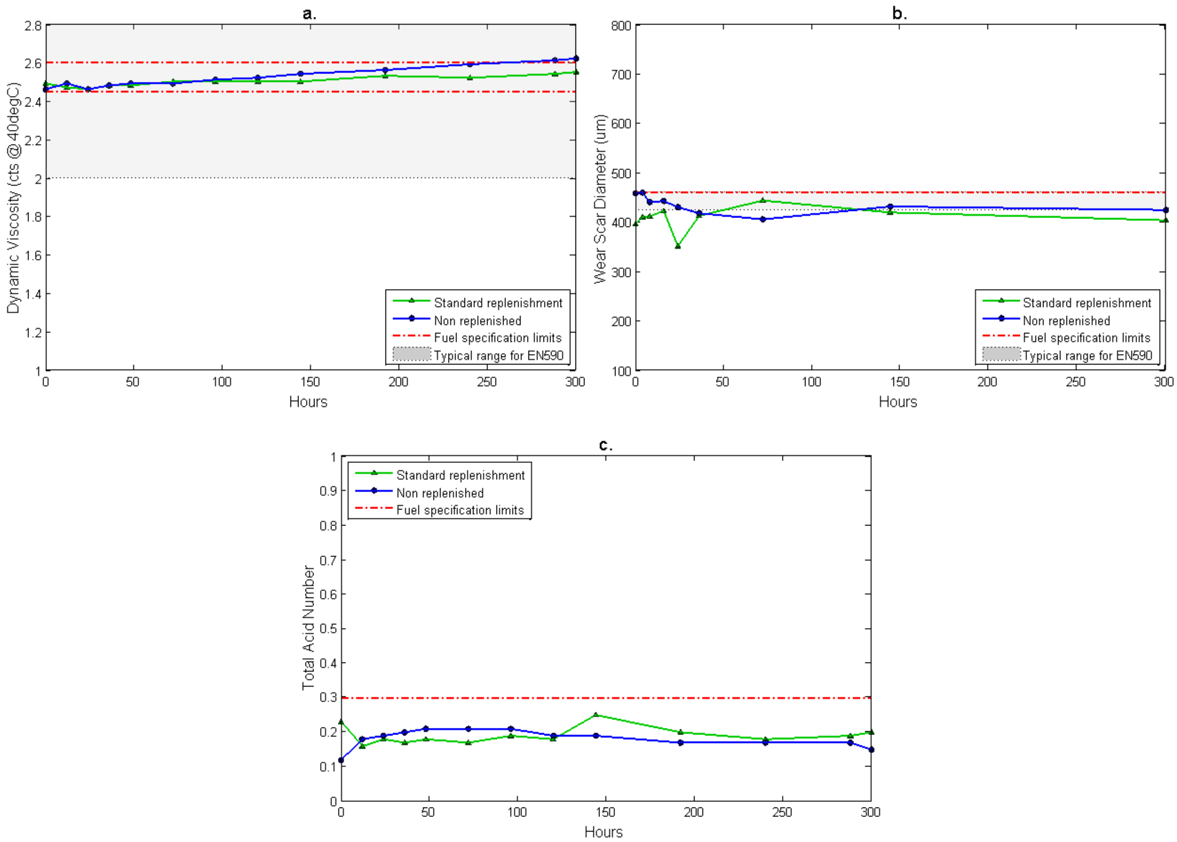


Figure 62: Physical and chemical degradation of ISO4113 fuel: a.) Dynamic Viscosity, b.) Lubricity, c.) Total Acid Number

Figure 62.a demonstrates the effects of fuel degradation on the observed viscosity of the ISO4113 fuel, showing that non-replenished fuel demonstrates an increasing viscosity through recirculation, while use of the standard replenishment schedule can be shown to reduce the observed rate of degradation. Figure 62.b then visualises the effects of recirculation on the observed lubricity of ISO4113, showing that even with replenishment, ISO4113 represents a more lubricious fuel than EN590 when recirculated. Figure 62.c then demonstrates that no significant indications of chemical degradation were observed in recirculation.

With respect to the GRTD fuel, the study concluded that the effects of fuel degradation through recirculation were undesirable, with evidence of significant and detrimental deviations from the desired hydro-mechanical properties of the fuel, regardless of the fuel replenishment schedule employed. Figure 63 shows the observed response of the representative hydro-mechanical and chemical properties of the fuel through the duration of the degradation study, for both use of the standard fuel replenishment schedule, and for the case of non-replenished fuel.

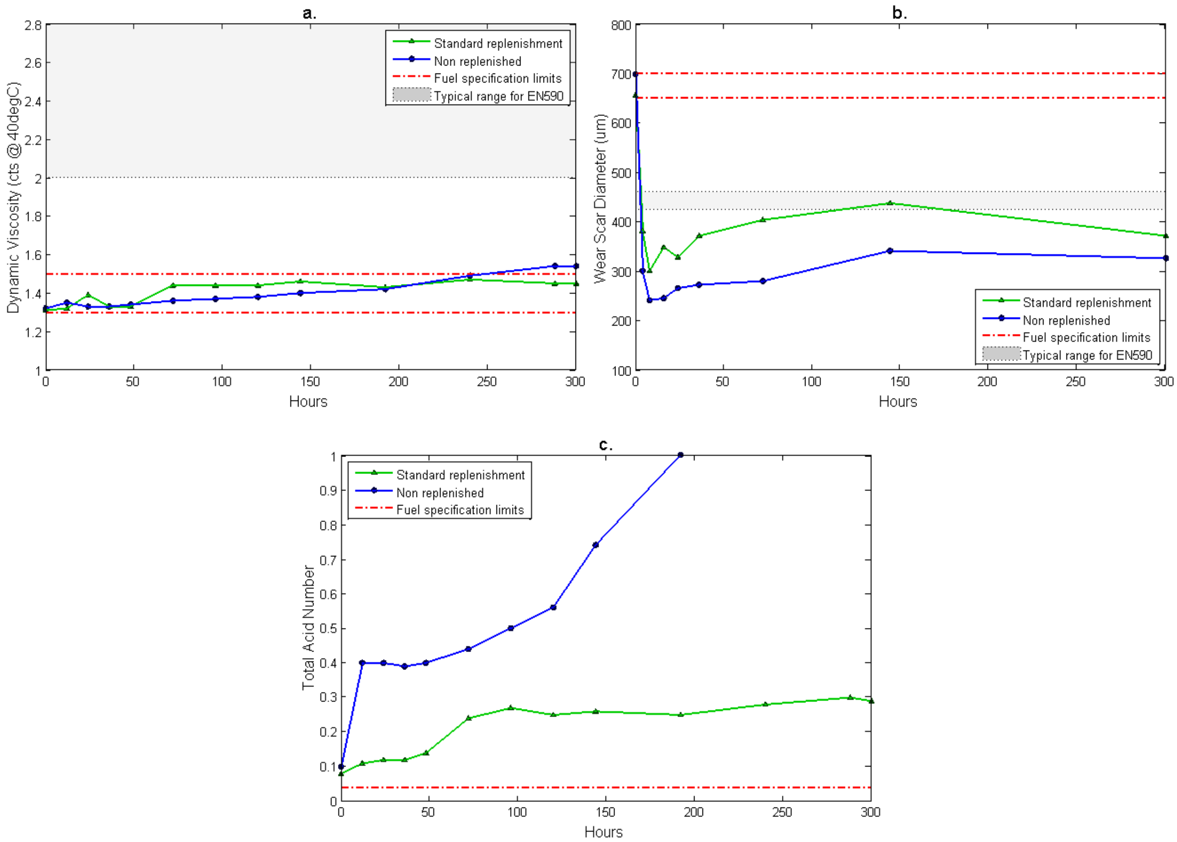


Figure 63: Physical and chemical degradation of the GRTD fuel: a.) Dynamic Viscosity, b.) Lubricity, c.) Total Acid Number

Figure 63.a shows the effects of recirculation on the viscosity of GRTD, demonstrating that a viscosity increase can be observed over time, and that replenishment can control the observed level of viscosity to be within the limits of the fuel specification. Figure 63.b then visualises the effects of recirculation on the lubricity of GRTD, showing that the fuel exhibits a significant change in its hydraulic properties within the first 10 hours of the test, representing a more lubricious fuel in recirculation, before then stabilising at a level influenced by the replenishment schedule. Figure 63.c then shows the effects of recirculation on the chemical degradation of GRTD, with significant chemical observed, stabilised by replenishment, but still corresponding to an increased propensity for IDID formation in recirculation. For both the replenished and non-replenished cases, the sample taken at zero hours can be seen to already exceed the fuel specification. The fuel stock used for this study had been used for a limited time in advance of this observation, so while it is possible that the fuel had been supplied outside of specification contrary to conformance certificates, it is also possible that the increased TAN observed is an indicator of chemical degradation of the fuel in storage.

With respect to the WWLTF fuel, the study concluded that the effects of fuel degradation through recirculation were less severe than with GRTD and could be controlled sufficiently through replenishment. Figure 64 visualises the observed response of the representative hydro-mechanical and chemical properties of the fuel through the duration of the degradation study, considering only the use of the standard fuel replenishment schedule as the only results available at the time of writing this thesis.

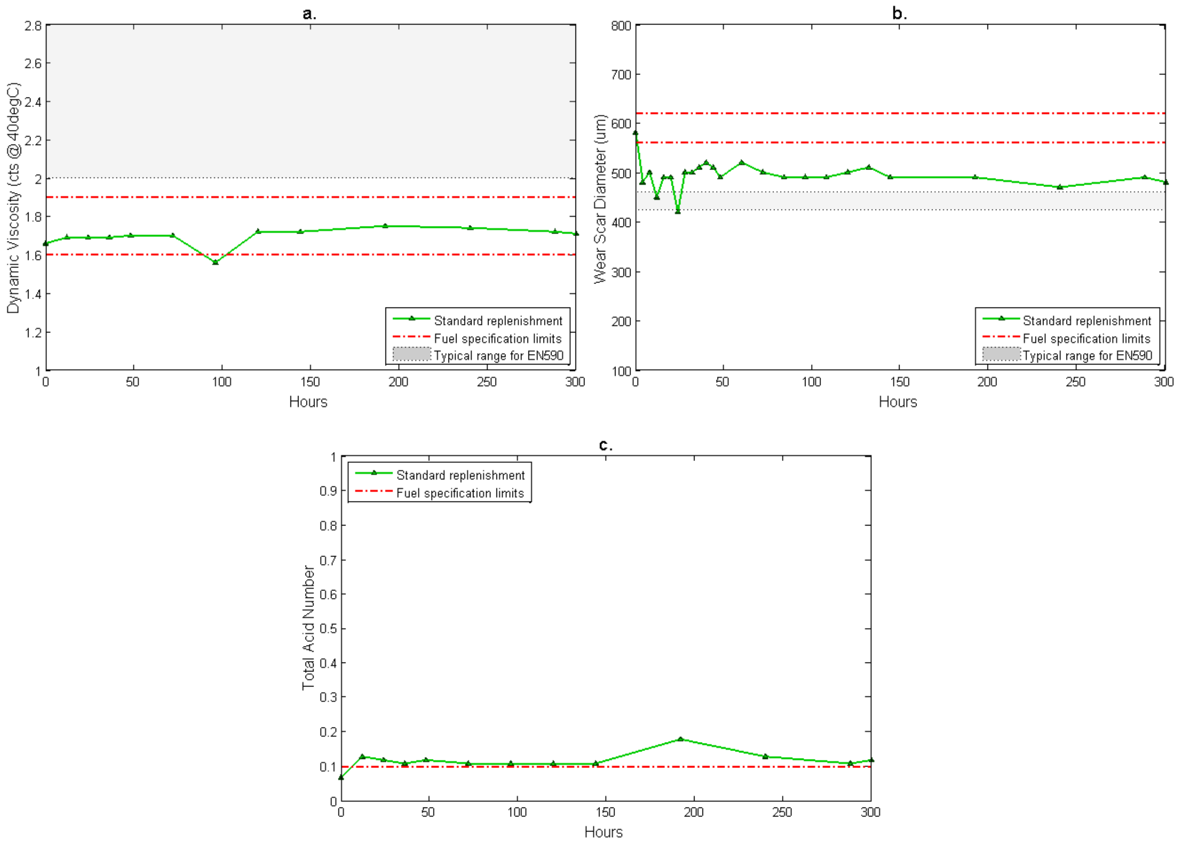


Figure 64: Physical and chemical degradation of the WWLTF fuel: a.) Dynamic Viscosity, b.) Lubricity, c.) Total Acid Number

Figure 64.a shows the effects of recirculation on the observed viscosity of WLLTF, demonstrating that the viscosity of the fuel generally remains within the limits of the fuel specification over time. Figure 64.b then visualises the effects of recirculation on the lubricity of WWLTF, showing that the lubricity degrades within the first 4 hours of usage before then stabilising at a level that represents a more lubricious fuel than specified. Finally, Figure 64.c visualises the observed chemical degradation of WWLTF in recirculation, showing that after an initial increase in TAN, the level of degradation is stabilised through fuel replenishment to a level just outside of the fuel specification.

Figure 65 compares the degradation in lubricity of WWLTF and both ISO4113 and GRTD. To provide a direct comparison, the observed degradation of each fuel is presented where the standard fuel replenishment schedule was employed. As can be seen, the degradation in lubricity of GRTD results in a fuel of equivalent effective lubricity to ISO4113 for recirculating tests. While WWTLF is shown to degrade such as to exceed the lower limit of its associated specification, in the case of recirculating tests, it still represents a less lubricious fuel than ISO4113 or EN590. This degradation in lubricity represents an improvement with respect to the performance of the equivalent GRTD fuel, but the level at which it stabilises at is still undesirable and may not present the required differentiation to the lubricity of the ISO4113 fuel.

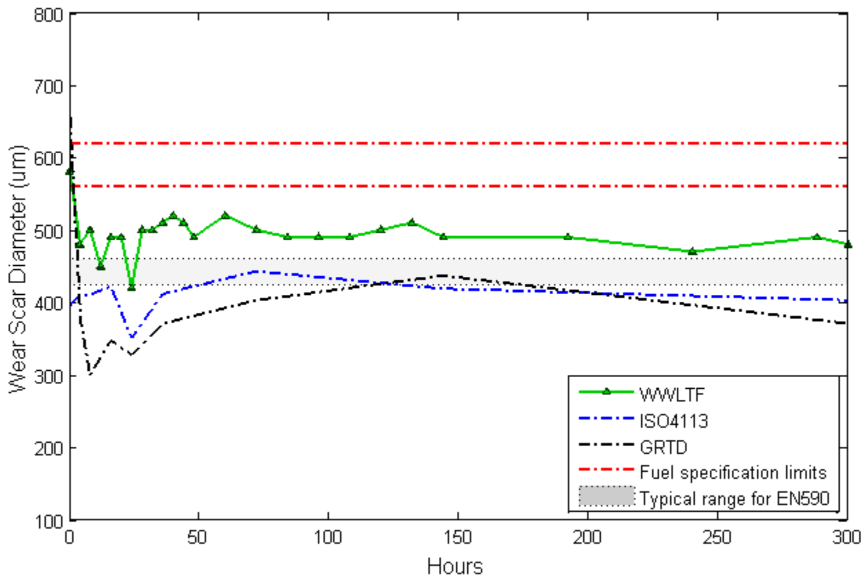


Figure 65: Comparative degradation in lubricity with standard replenishment schedule, with WWLTF specification limits shown

Figure 66 compare the observed chemical degradation associated with WWLTF, ISO4113, and GRTD. The chemical degradation exhibited by GRTD has been identified as a factor that potentially influenced the IDID formation that previously interrupted test progress. As shown, WWLTF exhibits a lower increase in TAN than GRTD, indicating that it represents a more chemically stable fuel than GRTD for recirculating tests, decreasing the propensity for IDID formation with its usage when compared to GRTD.

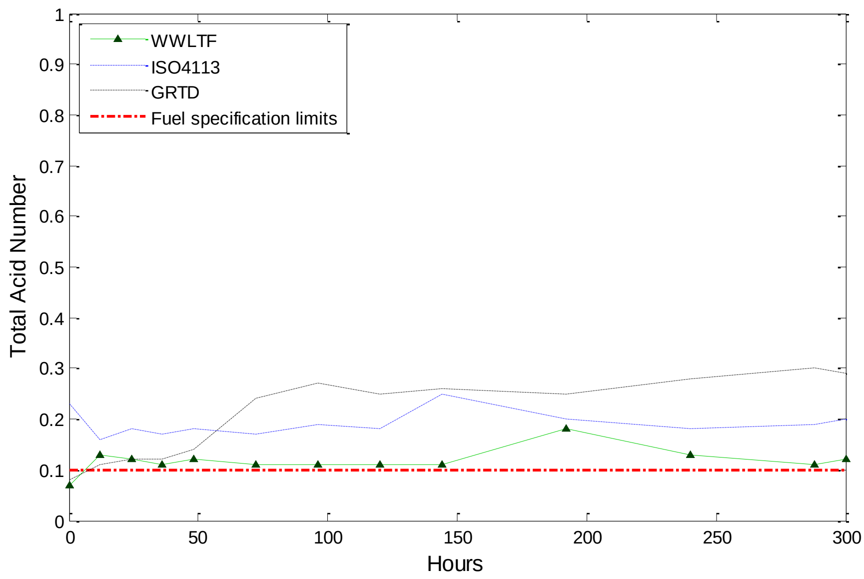


Figure 66: Comparative chemical degradation with standard replenishment schedule, with WWLTF specification limits shown

7.3.5 Learning from iterative experiments

Drawing on the lessons learnt from the preceding studies, the focus of the experimental investigation returned to characterisation of NCV bottom seat wear. The selection of design variables and levels thereof used in the first iteration of the design of experiments programme will be revisited based on the knowledge gained through the intervening experimental work. Furthermore, evolving business needs meant that any further experimentation would be taking place in a more resource sensitive environment, with the available testing resources limited in light of the demands of customer projects, and where the extended availability of the test resources was uncertain.

Through the results of the first iteration of the design of experiments programme, and subsequent investigations and characterisation of both IDID formation and IIT, it was decided that fuel inlet temperature would not represent a suitable design factor. While the fuel inlet temperature was identified as representing the environmental conditions associated with a given application, the results of the IDID investigation demonstrate that it would not be possible to meaningfully vary fuel inlet temperature whilst testing at IITs associated with robustness to IDID formation.

With a significantly reduced number of tests associated with the second iteration of the design of experiments, the decision was made to not run different treatment combinations in parallel on a double headed rig as previously detailed in §7.3.4, but instead to increase the sample size for each test, resulting in a maximal of 12 repeated measurements for each treatment.

7.3.6 Choice of methodology

Of the possible Experimental Design methodologies discussed in §4.4, both Factorial Designs and the Taguchi Method were considered for application to this experimental study. The TM was considered as it would allow for resource efficient design, allowing a high number of variables to be assessed at multiple levels, allowing main effects and their linearity to be studied. However, the downfalls of the TM, as discussed in §4.4.3, were significant with respect to this experimental programme, both with respect to the confounding of main effects and interactions, and also with respect to augmenting any initial experimentation through projection or addition.

As such, the classical factorial design methodology was selected, specifically fractional designs of a resolution that would allow for main effects and interactions to be assessed without confounding. A fractional design would also best allow for further study, either through increasing the resolution into a full factorial design, through projection for additional variables, or though the addition of centre and/or axial points. With a view to the iterative nature of the experimental process, FFD was identified as the most suitable methodology.

7.3.7 Selecting factors and their levels

Table 15 presents a summary of the possible usage variables discussed in §7.3.2. & §7.3.3, and their suitability for inclusion in this experimental study. Their suitability is also discussed in terms of the knowledge derived in the iterative experimentation detailed in §7.3.4.

Group	Variable name	Suitability for this experimental study
Usage variables where consensus was met of being of high significant to NCV seat wear	Number of injections	Yes – in combination with injection durations
	Fuel lubricity	Yes – in combination with fuel viscosity
	Hard particle contamination	Yes
	Rail pressure	Yes
Usage variables where informed disagreement was met of being of high significant to NCV seat wear	Oil-in-fuel dilution	No
	NCV pin velocity	No – dependant variable only
	Fuel viscosity	In part – as a dependant variable in combination with fuel lubricity
	Environmental temperatures	In part – interaction with rail pressure can result in IDID
	Injection durations	In part – as a dependant variable in combination with number of injections
	Air ingress	No
	Pressure dilation	In part – as a dependant variable with rail pressure

Table 15: Summary of variables with respect to this empirical study

For the purpose of this experimental study, and in order to best replicate the fuel conditions associated with application, a controlled level of hard particle debris contamination will be employed such as to be equivalent to the total number of hard particles in usage, but not representing an acceleration over typical application conditions. As such, hard particle debris contamination would not be considered as a design variable in this study.

The usage variables selected to be used as design variables in this experimental study were therefore system operating pressure, the use of advanced injection strategies in the form of multiple injections, and fuel type.

As discussed in §3.4 system operating pressures are a product characteristic that has been the focus of significant development. Maximal system operating pressures have increased greatly for DPEs Euro VI family of products when compared to previous systems, while the common rail technology employed results in a higher average hydraulic pressure. At the same time, advancements in high efficiency SCR systems, and associated re-optimisation of combustion, have seen some OEMs splitting their vehicle strategy for Euro VI dependant on application. As such, some vehicles utilising Delphi Technologies Euro VI fuel systems are doing so using the maximal system operating pressures (2400-2700bar) while others are using operating pressures more akin to Euro III/IV/V systems (1800-2200bar). As such, it was decided that the two pressure levels to be used in this experimental study should be representative of those two tiers, while remaining within the operating range of the FIS used, and the low and high levels were selected to be 1800bar and 2500bar respectively.

As detailed in §3.4, advanced combustion strategies enabled by the used of multiple injection strategies, can allow the OEM to reduce the emissions and fuel consumption of a vehicle, and are partly dependant on EATS strategy. As such, in comparison to previous systems, there is an increased adoption of smaller pilot and post injections for many Euro VI systems. It was therefore decided that the low and high levels of this design variable should be a single main injection event only, and a typical pilot-main-post coupling of 3 injections. To reduce the potential influence of fuel flows, both on flow erosion and the total number of hard particles through the system, the total injected quantity would be kept the same regardless of the number of injections. In order to improve the efficiency of the test with respect to the number of injections over a given test duration, the tests were run at a constant speed that is typical of a maximum rated speed for the application of this given system. Furthermore, the use of non-combusting hydraulic rigs allows injection events to take place every engine crankshaft revolution, as opposed to the every-other revolution typical of a 4 –stroke engine application.

The low and high levels of fuel types to be used in this experimental study were to be chosen based primarily on their lubricity value, with acknowledgement to their viscosity as discussed in §7.2. In order to best represent the typical Class 1 fuel with ‘good’ lubricity, ISO4113 calibration fluid would be used in lieu of EN590 diesel, while to best represent the Class 3 fuels associated with some of the emerging markets for Euro VI fuel systems, the fuel chosen to represent ‘poor’ lubricity was WWLTF.

7.3.8 Definition of selected experimental design

For the reasons discussed in this section, the experimental design represented a 2-level, 3 factor, half fractional factorial design, represented in the factorial form by 2^{3-1}.

This represented a total of 4 treatment combinations that were to be tested, in series, and in a random order, on the available double headed test rig. Each treatment combination would utilise a total of 12 samples, offering a form of pseudo-replication as discussed. Each test would be conducted for 1000 hours, with performance characterisation testing completed at 0, 500, and 1000 hours, while the test samples would only be disassembled and measured after 1000 hours had been completed, as described in Table 16.

Test Feature	**Count, duration, or time interval**
Number of tests	4
Number of samples per test *(including 1x electronically disconnected sample per test)*	12
Duration of each test	1000 Hours
Performance characterisation intervals for each sample and test	0 Hours 500 Hours 1000 Hours
Disassembly and component measurement interval for each sample and test	1000 Hours

Table 16: Test count, sample allocation, duration, and measurement intervals

Table 17 summarises the experimental design, while Figure 67 provides a visualisation in experimental space.

Variable Name	**Low Level**	**High Level**	**Variable type**
Fuel type (lubricity)	ISO4113	WWLTF	Discrete
System operating pressure	1800	2500	Continuous
Number of Injections	2	6	Continuous

Table 17: Design variables and levels

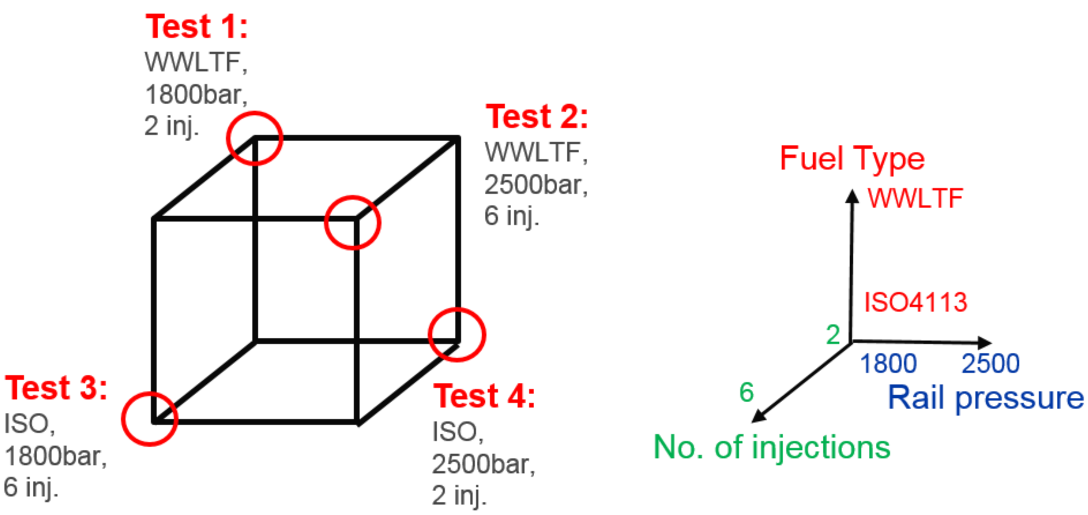


Figure 67: Visual representation of the experimental design

7.4 Physical experiment set up

7.4.1 RPRCS system

The FIE system to be used for this experimental study is one of the Euro VI F2 family of products, known as F2R. The F2R system features a remote, 2 or 3 plunger, engine oil lubricated pump with integral low-pressure transfer pump, rather than the distributed pumps used elsewhere in the F2 family of products. The F2R system is rated to 2500bar system operating pressure, and is designed for medium and heavy-duty applications, with engine cylinder capacities in the 1 to 2.6 litre range and is available as both 4- and 6-cylinder systems.

For this experimental study, the system will be used in a 3-plunger and 6 injection configuration, and will therefore comprise of: a 3-plunger high pressure pump, a common rail, 6 injectors, 7 high pressure pipes, and a Delphi Technologies ETC. A 2700bar pressure sensor is mounted to the common rail and rail pressure control is achieved primarily through an electronically controlled Inlet Metering Valve (IMV) integrated into the high-pressure pump. Furthermore, an electronically controlled Pressure Control Valve (PCV) is fitted to the common rail, which can add additional functionality with respect to rail pressure control but is primarily used to rapidly drop rail pressure in the event of the engine being switched off, or as a safety feature in the event of overpressure events.

This F2R system utilises the DFI5 injector, which is typical of the Euro VI F2 family of products as described in §3.2.2.

7.4.2 Durability rigs

The test rig to be used in this experimental study was designed to be capable of running the equivalent of two engine fuel injection systems in parallel. An electronic motor is used to provide highly controllable drive to the fuel pumps, enabling equivalent speed and load changes as would be typical of an engine. The drive motor is controlled by an automated test plan, itself highly customisable in terms of equivalent engine speed, and load demanded of the FIS. A software system, ‘VPRS’ is used to control and monitor the test rig autonomously. In combination with a CO2 fire suppression system, the automated test plan can allow for 24 hour a day running, with monitoring, alarming, or controlled shutdown based on a customisable number of measurements taken from the rig, and conditions monitored through the ECUs.

7.4.3 Low pressure fuel system

For testing on hydraulic rigs with controlled hard particle contamination, a recirculating system using fuel tank integrated into the test rig is used. In order to enable the controlled level of hard particle contamination, additional hardware and electronic monitoring and control are required. A simplified visualisation of the system is shown in Figure 68.

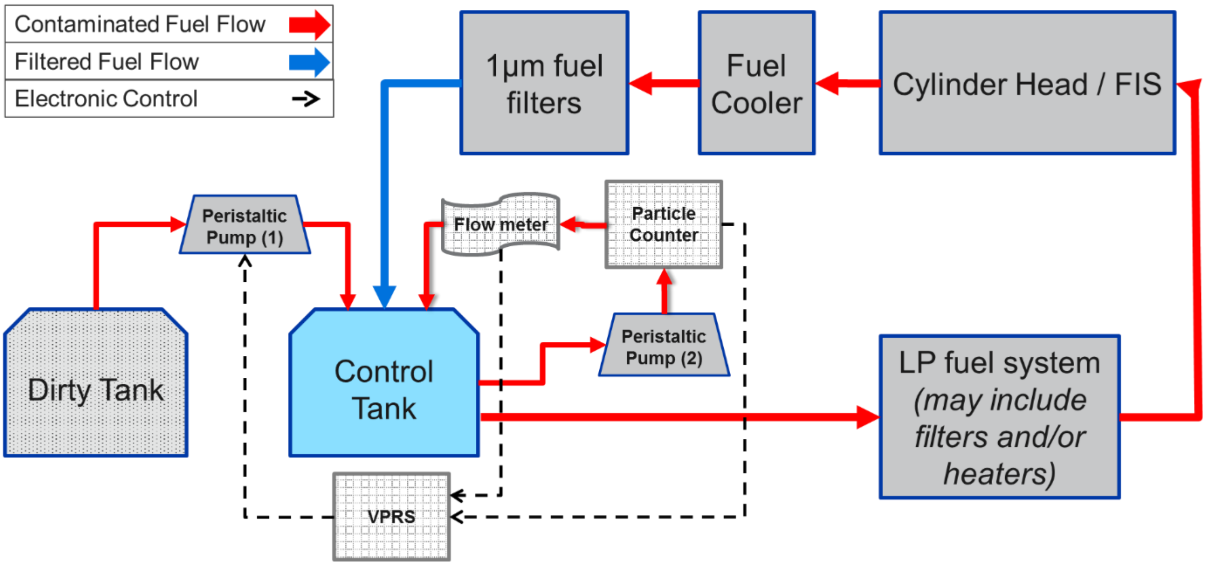


Figure 68: simplified visualisation of low-pressure fuel system for testing with hard particle contamination

In such a system, the main rig fuel tank is bypassed completely, instead using two parallel fuel tanks, one designated as the 'dirty' tank, and the other as the 'control' tank. At the start of a test, the control tank is filled only with fresh fuel, free of any hard particle contaminants, while the dirty tank is filled with a controlled combination of fresh fuel, and a mass of hard particle contaminant. To ensure that any hard particles remain suspended within the fuel and do not settle at the base of each tank, constant agitation is supplied by spindles that remain driven when the test rig is idle. A peristaltic pump ((1) on Figure 68), is used to flow fuel taken from the control tank through a Klotz particle counter and a flow meter in series, both of which are monitored by the VPRS system.

The Klotz particle counter requires a controlled flow level in order to optically count, and determine the size distribution of hard particles, providing an output of particle per ml in the 3 different size ranges that correspond to the ISO 11171 classification (<4um, <6um and <12um). The VPRS software monitors the flow meter to ensure that the flow through the particle counter is within the required range, and then typically uses the particle count in the 6um range as a control variable. The observed cleanliness of the control tank, in terms of that 6um range, is then used by VPRS to control the speed to a separate peristaltic pump ((2) on Figure 68) to supply contaminated fuel from the dirty tank in order to maintain, or increase, the level of contamination in the control tank as required by the specified contamination level.

When the control tank is controlled at the required level, fuel flow is supplied to the FIS system, via a LP transfer pump. The LP system may include filtration in order to demonstrate the effectiveness of application filtration or fuel heaters to replicate higher operating temperatures. For the purposes of this thesis, no filtration or heaters were employed in the LP system. After the controlled fuel has been through the HP FIS system as either injected or otherwise returned, it is firstly cooled, before then passing through a filtration system optimised to remove any hard particle contamination. Doing so allows for fuel to recirculated in order to reduce resource requirements for a given test, while best enabling a controlled level of hard particle contamination to be maintained.

For the purposes of this experiment, the contamination used is the Arizona A1 test dust, matching the particle size and hardness distribution typically associated with application.

7.4.4 Data Capture

The measurement capturing process available on the durability rigs can be described as consisting of 4 layers. The lowest resolution system is that of discrete measurements that require user intervention. A low frequency data measurement system is then controlled through the test rig software and is specified by the responsible engineers. In parallel, a high frequency data measurement system is also controlled by the test rig software and can be specified by the responsible engineers. Finally, measurements can be calculated through post processing of any of the proceeding layers.

The user intervention measurement systems typically take the form of measurements or samples taken at discrete intervals as specified by the responsible engineer, such as fuel samples taken from the rig tanks for subsequent analysis that would not be possible using on-rig instrumentation.

The test rig software controlled low frequency data measurement system is typically used to monitor system operation through the course of an endurance test. The user can specify a number of channels to monitor and can define the measurement capture frequency. This system is typically used to capture system operating pressures and temperatures to monitor system performance, and to monitor electronic test control parameters. Furthermore, the environmental conditions, such as the ambient temperature of the test cell and the fuel inlet temperature, can be monitored for control purposes, and to ensure that specified test conditions are being met. This system can also be used to monitor ETC channels, including both the demanded and observed values of rail pressure, and currents supplied to electronic control valves.

In addition to recording the conditions over the duration of the test, the low frequency data measurement system is also used to provide automated control over key parameters as specified by the responsible engineer. The engineers can specify lower and/or upper limits for critical parameters such that the rig will automatically enter a controlled shutdown process in the event of breaches of those limits. Furthermore, for times when Test Rig Support personal are available to attend the rigs, channels with alarms on are displayed on the rig control panel, with warnings for when those channels are approaching their limits.

The high-speed data measurement system is typically used to monitor the electrical signals provided to the FIE system, and the signal provided by the integrated rail pressure sensor. Current and/or voltage clamps are used to provide an unobtrusive measurement of the electrical signals provided by the ECU to the FIE system, allowing the engineer to view and manipulate high frequency waveforms. With regards to the high-pressure pump, this allows the engineer to assess the metering valve position through an engine cycle, with the possibility of post processing to assess changes in system performance. Similarly, the drive waveforms supplied to the NCV can be used to infer a number of characteristics of the injector, including electromagnetic flux and stator resistance. The rail pressure sensor signal returned from the ETC also be captured in high frequency, allowing for further interrogation of the system performance through an engine cycle, and over the course of a test.

The final layer of data measurement on the durability rigs is a post processing layer, typically used to infer additional parameters from the available low and high frequency data. A typical example is the calculation of the stator core temperature through its resistance value, determined through a combination of known values at an ambient temperature, and voltage and current waveforms captured at high frequency during the test. Techniques have also been trialled to infer fuelling changes over time using high frequency data for the rail pressure sensor and metering valves, but require additional development using empirically derived models.

7.4.5 Performance rig

To best ensure research quality with respect to both repeatability and validity, all injector performance characterisation tests associated with this research were conducted using the performance rig described in §5.3 in an unmodified form, with trained personnel following standard operating procedures, and using the standardised, and controlled test plans.

7.4.6 Test samples

To best ensure research repeatability and validity, all samples used where standard fuel injectors typical of high-volume manufacturing and assembling processes. For each test, 12 samples were to be used, as chosen from a larger sample population at random. For each treatment combination, the response metrics could then be expressed in terms of both location and dispersion, making it possible for sample means to be compared in terms of confidence bounds. Furthermore, statistical outliers could also be identified.

In order to provide a form of experimental control, while retaining statistically significant sample sizes, one injector from each set of 12 was used as an electronically disconnected control injector. These samples were installed into the durability rig such that they remained part of the hydraulic system, but where never electronically actuated, meaning they were still exposed to the same system operating pressure and temperatures, but were not subjected to any injection events through the DoE tests. These control samples would be subject to the same measurements, in the form of performance testing and physical component measurements as the remaining test samples.

7.5 Performance Metric Development & Selection

While a number of existing metrics for injector performance characterisation were in common usage within Delphi Technologies as detailed in §5.3, in order to best characterise the relationship between the design variables, injector performance, and the physical wear to the NCV, some development of both existing and new injector performance metrics was required.

7.5.1 Maximal Drift

The most common means of expressing injector performance drift within Delphi Technologies was through the observation of maximal drift. The maximal drift was typically determined for either an entire gain curve, or zones within that gain curve, and represented the upper extent of the fuelling change observed, expressed as either a fuel quantity (in mm^3 or mg), or as a percentage of the as-new performance of the injector. It is possible to express maximal drift for an injector in the untrimmed condition (UM), which would include any effects of timing changes, the trimmed condition (TM), or for both conditions.

As discussed in §5.3.2, the reporting of fuelling drift varied between customer projects. With no specific customer reporting required for this experimental investigation, the decision was made to consider the ballistic region of the fuel gain curve as a single fuelling ‘zone’ that represented injected fuels between $5mm^3$ and $40mm^3$. As $5mm^3$ injections typically represent the lowest injection quantity used in application, this 5-$40mm^3$ zone represents an appropriate range of the ballistic fuelling curve while being within the capability range of the FMU used.

Furthermore, the form of the observation, and the means through which it was derived varied between customer project teams, and to an extent, between individuals within those teams. Some individuals would interrogate the fuelling data to calculate the fuelling change, whilst others would rely on reading the numbers from the graphs generated through Matlab or Excel.

A more formalised, consistent, and rigorous approach for reporting drift would be required for this experimental programme to ensure observations were made in an accurate, repeatable, and traceable manner. In order to achieve this, a Matlab function as developed to manipulate the fuelling gain curves that corresponded to multiple injector performance characterisation tests over time such that the maximal fuelling change observed in each defined zone of the gain curve. The function is capable of reporting both UM and TM for a population of injector samples in a single execution, not only ensuring a consistent approach to Drift calculation, but also representing a more resource effective means for doing so. Furthermore, the function can then overlay the observations onto the typical visualisation of the fuelling change for an injector, or sample of injectors, creating an efficient and rigorous means for reporting fuelling changes. An example of this visualisation is shown in Figure 69.

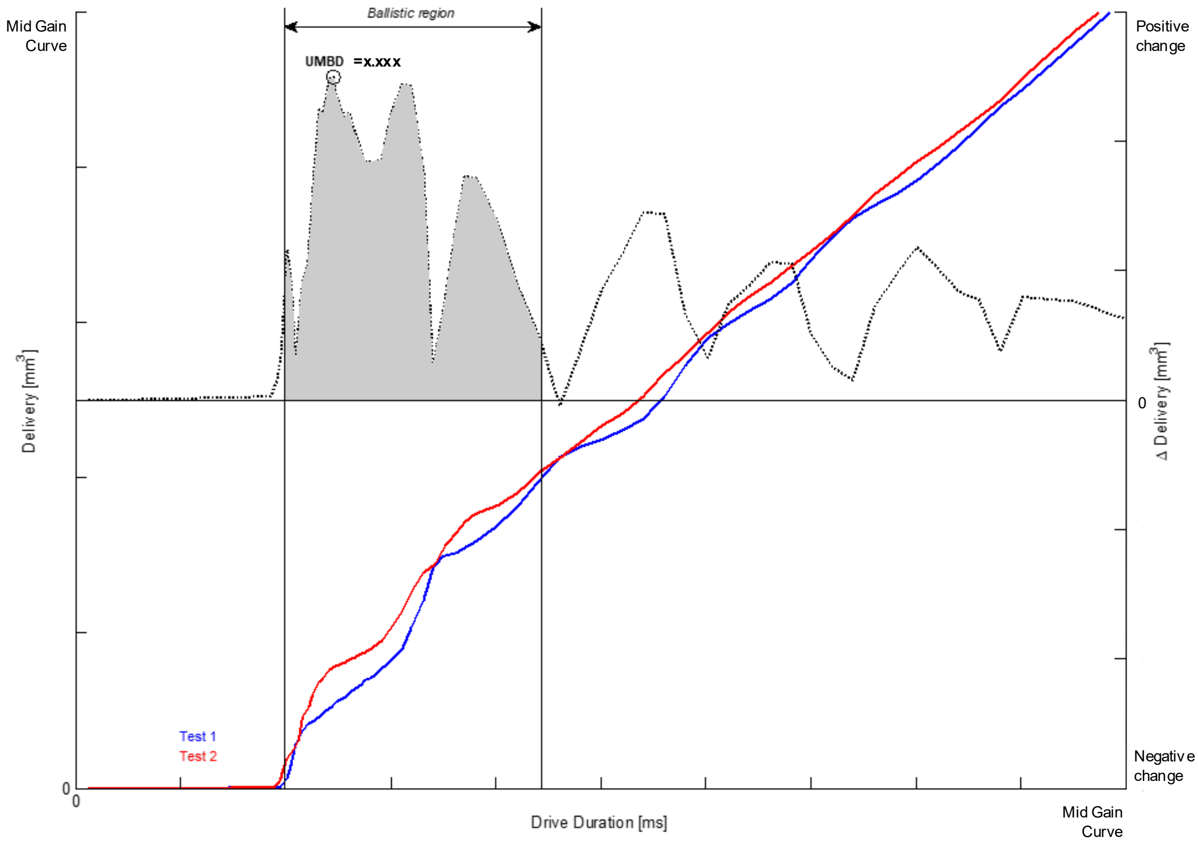


Figure 69: Visualisation of fuelling change between two tests, with maximal drift (UMBD) annotated

7.5.2 Integral of Drift

Calculation of the integral of an injector's fuelling change is a new metric associated with this experimental study. After observing that individual injectors could present relatively high maximal drift in highly localised areas of the fuelling gain curve only, an alternative form of representing an injector's drift was sought. Around the points of inflection in the ballistic region of the gain curve, it is possible for an injector to exhibit relatively high maximal drift, while otherwise demonstrating low fuelling change. It was proposed that taking an Integral of Drift within a zone of the fuel gain curve would provide a means for differentiating between such injectors that exhibit a change around the points of inflection only. Figure 70 provides a visualisation of two injectors that exhibit similar magnitudes of Maximal Drift, expressed as UM, in the ballistic region of the gain curve, as evidenced by the annotation identifying their respective fuelling changes. An additional layer of annotation then identifies their respective Integral of Drifts, expressed as an Untrimmed Integral (UI), with the area under the fuelling change curve shaded grey. As can be seen, the injector in Figure 70.a demonstrates much greater fuelling change throughout the ballistic region when compared to the injector in Figure 70.b. which exhibits a slightly higher maximal fuelling change around the points of inflection only. This differentiation is only present when the Integral of Drift metrics are compared.

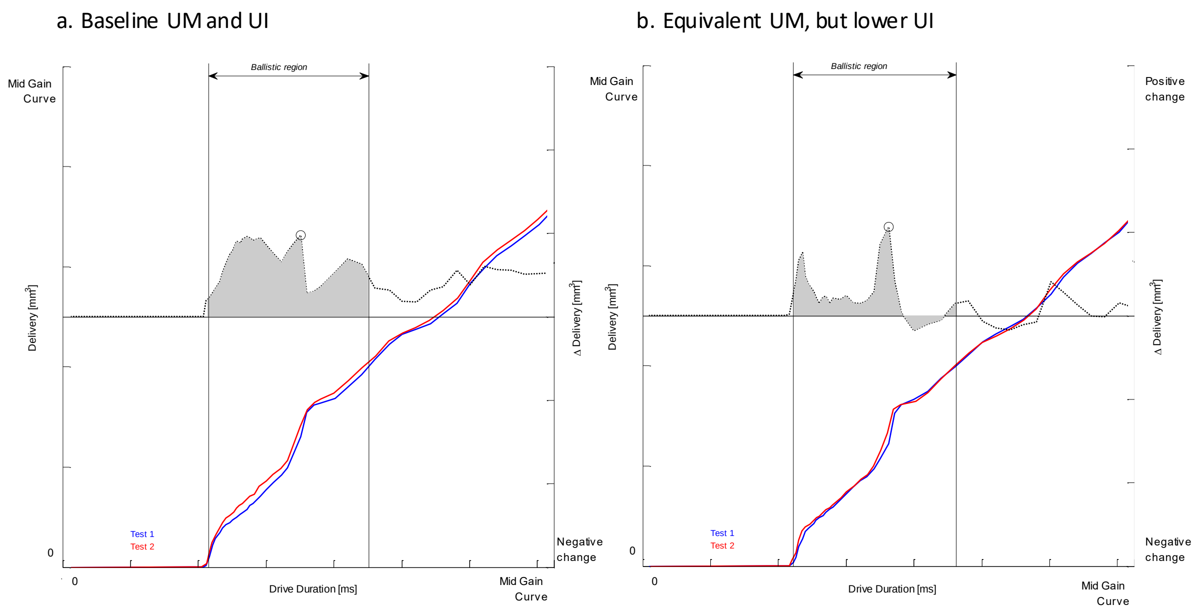


Figure 70: Visual comparison of two injectors that show equivalent drift expressed as UM, but different levels using the UI metric

In order to calculate the Integral of Drift, a further Matlab function was developed. The function can determine the Integral of Drift for injectors in both the untrimmed (UI), and trimmed condition (TI), and can handle multiple injectors in one execution. After determining the fuelling difference curve for the multiple tests of one injector, the function then integrates that curve with respect to injection logic. In the event of the use of zones of the gain curve defined in terms of fuelling, the function determines the injection logic range that corresponds to the fuelling range for the injector in the as new condition.

7.5.3 Performance metric selection

For the purposes of this research, a decision was made to use both the existing, UM drift metric, expressed as a change in injected fuel volume (mm^3), and the newly developed TI drift metric, expressed as an area between two fuel gain curves (mm^3*us). By continuing to use the UM drift metric, comparisons could be made to previous results in a form that is already reported and understood by the engineering community of Delphi Technologies. However, in also using TI drift, a direct and meaningful comparison will be made available with respect to how both metrics respond to varying levels of usage variables, and how they infer NCV seat wear. In addition to UM and TI drift, MDP timing change, expressed in microseconds (µs), will also be reported for each treatment combination, representing a direct measurement of any effects on the injector timing only.

As such, while UM drift represents the combined effects of timing and shape changes to the gain curve, TI drift isolates any effects on gain curve shape only, while MDP timing change isolates any effects on injector timing.

7.6 Physical wear metrics

In order to measure NCV bottom seat wear, both in terms of wear to the NCV pin bottom seat and the corresponding surface on the PG top face, a Taylor-Hobson Talysurf machine was used. While alternative systems, including optical systems, are available and in mixed used throughout Delphi Technologies, the physical trace associated with the Talysurf represents the most widely used measurement, allowing for simplified post processing, and direct comparison with previous results.

For every injector sample, a total of 8 measurement traces will be taken. For the NCV bottom seat, 4 traces will be taken that correspond to 90° intervals of the pin circumference, while 4 corresponding traces will be taken on the PG top face.

7.6.1 NCV pin seat wear

Figure 71 provides a simplified visualisation of the method used for determining the magnitude of NCV pin bottom seat wear. The scales used for the x and y axes are such as to exaggerate the profile of the seat and any associated wear to make observations easier. As can be seen, the measured trace is compared to both the nominal profile from the drawing, with no tolerances visualised, and the extrapolated as manufactured profile. The difference observed between the as manufactured profile, and the as worn profile, is then used to determine the magnitude of wear to the pin bottom seat. By taking an average of the 4 traces, this method has been proven to be an accurate and repeatable means for measuring wear to the NCV pin bottom seat, expressed in microns (µm).

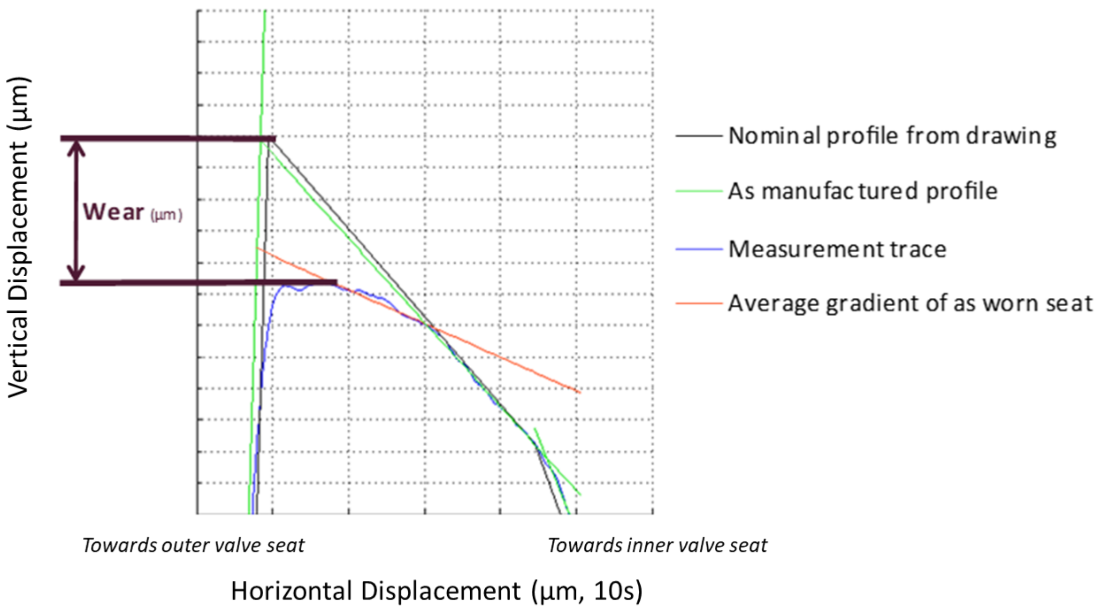


Figure 71: Simplified visualisation of method for determination of NCV pin wear

7.6.2 PG trench depth

Figure 72 provides a simplified visualisation of the method used for determining the magnitude of wear to the seat on the PG top face. The scales used for the x and y axes are such as to significantly exaggerate the profile of the wear to facilitate observations. The measured traced for the trench worn into the PG top face is compared to the unworn surface outside of the NCV pin seating diameter. In a similar way to the NCV pin bottom seat, in taking an average of the magnitude of wear from the four traces, this method has been proven as an accurate and repeatable means for measuring PG trench depth, express in microns (µm).

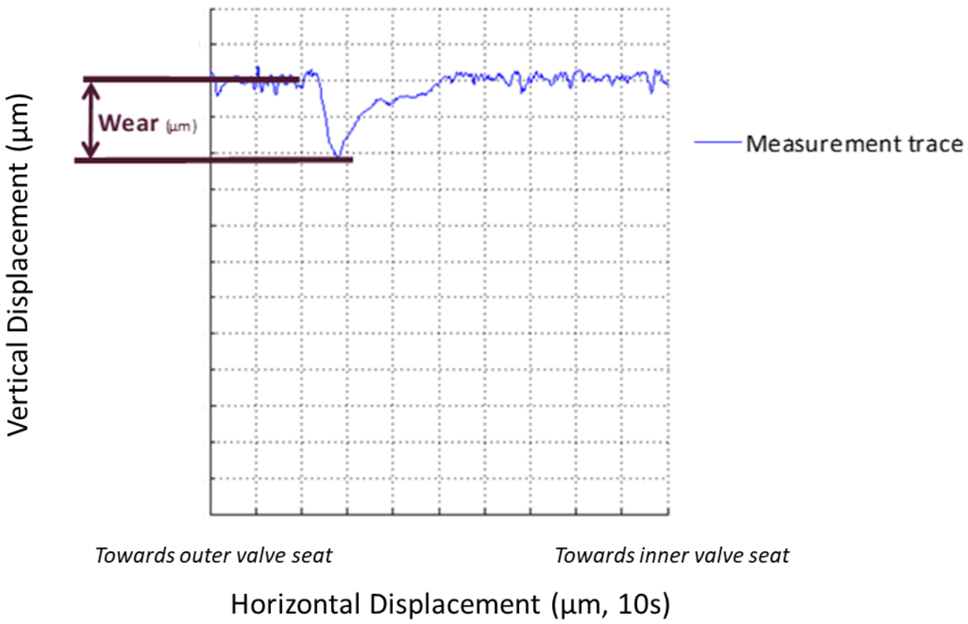


Figure 72: Simplified visualisation of method for determination of PG trench depth

7.7 DoE results – Performance and Measurement Metrics

The results for each metric, both those related to performance and the physical wear of the components will be expressed in a standardised format for each of the tests. For each metric, a 4-way visualisation of the results will be presented in terms of change from observation at zero hours, which for the performance metrics, will include the observations from both the 500-hour performance characterisation test, and the 1000-hour performance characterisation test. For each metric, the results will be presented in terms of the individual values of the observations for each test, a boxplot of the distribution of those values including the identification of any outliers, a plot of the sample mean of each test with 95% confidence intervals identified, and a visualisation of the relationship between the mean for each test and test duration.

The individual value plots demonstrate the most basic visualisation of the data for each treatment combination, visualising the individual observations associated with each sample on test. As shown in Figure 73, for each treatment combination, the observations associated with both the 500- and 1000-hour performance characterisation tests are shown, with colours used as the differentiator. Furthermore, the observations associated with the electronically disconnect control injectors are visualised alongside the test samples, using a different symbol as the differentiator.

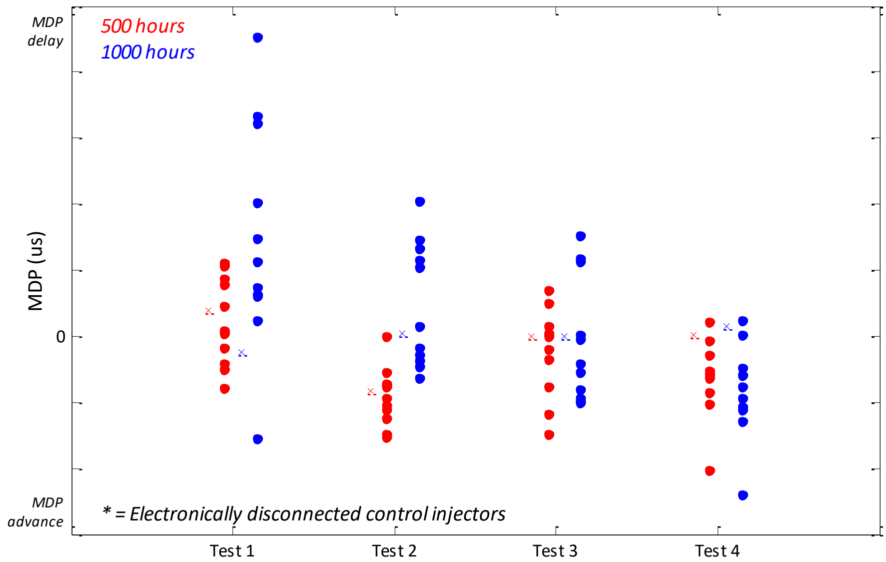


Figure 73: Individual values for each treatment combination, for both the 500- and 1000-hour intervals

The boxplots then provide a visualisation of the statistical distribution of the observations associated with each treatment combination, an example of which is shown in Figure 74. The boxplots visualise the median, range, and quartiles of the distribution, while identifying population outliers that are greater than 1.5 times the interquartile range either above or below the third and first quartiles respectively, visually differentiated as single '+' points outside of the boxplot.

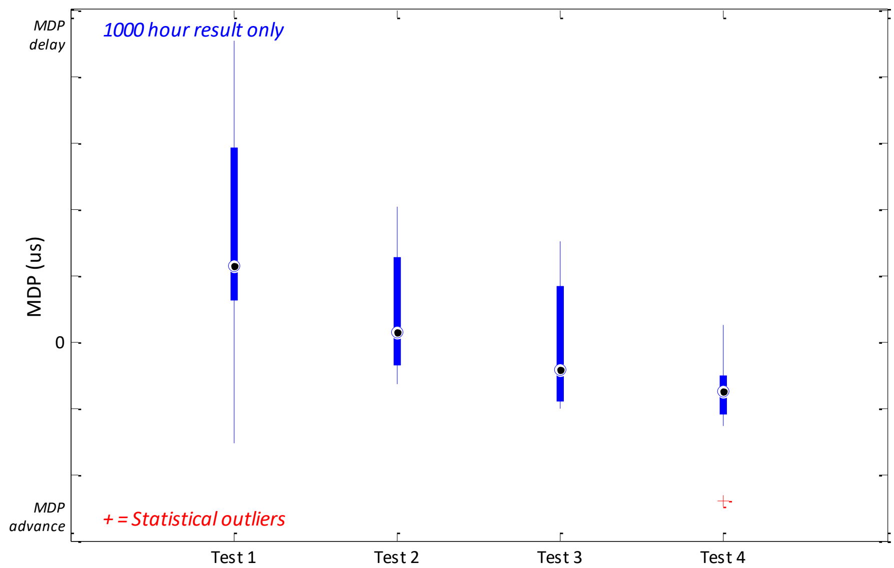


Figure 74: Box plot for 1000-hour results, with outliers identified (>1.5xIQR)

The visualisation of the mean with 95% confidence interval then provides a visualisation of the location of the response for each treatment combination, an example of which is shown in Figure 75. The sample mean for each treatment combination is calculated excluding any outliers previously identified. The 95% confidence intervals are then calculated from the sample variance, with a correction applied using the t-test based on the sample size, which will differ between treatment combinations dependant on the number of outliers excluded. The width of the confidence interval therefore provides a visualisation of the dispersion associated with each treatment combination.

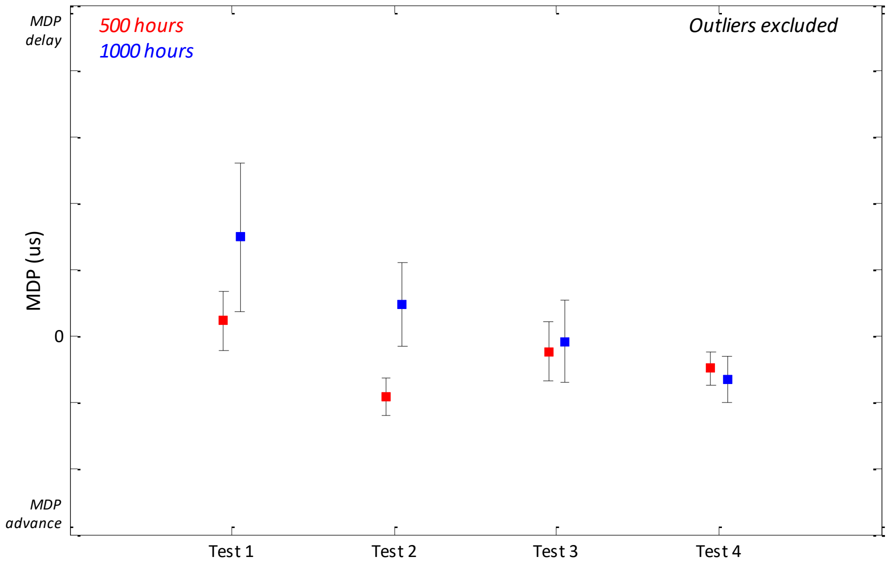


Figure 75: Sample means for 1000-hour results with outliers excluded, and 95% confidence intervals shown

The final plot provides a visualisation of the progression of each metric over time for the different treatment combinations, and example of which is shown in Figure 76. This plot allows for observations to be made about the relationship between each performance metric and time, allowing inferences regarding linearity and regression over time. Each treatment combination is differentiated by colour, with sample means and 95% confidence intervals for each observation. The two bogey tests are then differentiated with dashed lines as appropriate. To facilitate comparison between the different treatment combinations, a horizontal offset is applied to each observation to avoid the sample means and confidence intervals from stacking over each other.

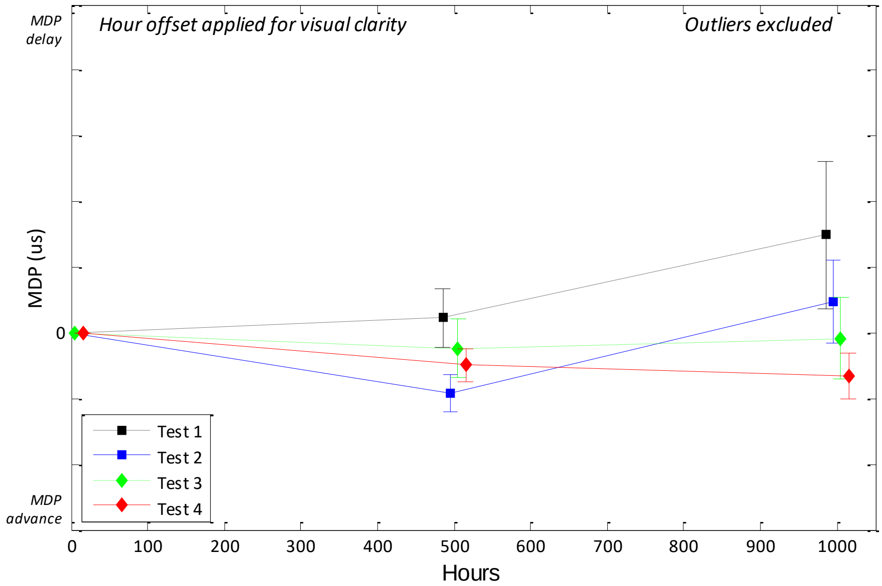


Figure 76: Progression of sample means and 95% confidence intervals over time

7.7.1 Minimum Drive Pulse

The results for the MDP metric for each treatment combination are presented in Figure 77. The visualisations of the samples means and 95% confidence intervals indicate a possible trend, where treatment combinations 1 & 2 demonstrate a positive MDP change, indicating a timing retardation, whereas combinations 3 & 4 demonstrate either no significant timing change, or a timing advance. Furthermore, the visualisation of progression over time shows that treatment combinations 1 & 2 exhibit a significant timing change between 500 & 1000 hours. Treatment combinations 1 & 2 both feature the WWLTF fuel, which as demonstrated in §7.3.4 is subject to a degree of chemical degradation that may increase the propensity for IDID formation. The observations associated with test 1 at 1000 hours also demonstrate more dispersion than those of the other tests. The visualisation of the progression of MDP timing change with time shows that treatment combinations 3 & 4 show a trend for negative timing change with time (an advance in valve timing), while treatment combinations 1 & 2 demonstrate a reversal in MDP timing change between 500 and 1000hours, ultimately exhibiting a positive timing change with time. The control samples can be observed to exhibit lower absolute timing change, but are typically within the range of the sample population.

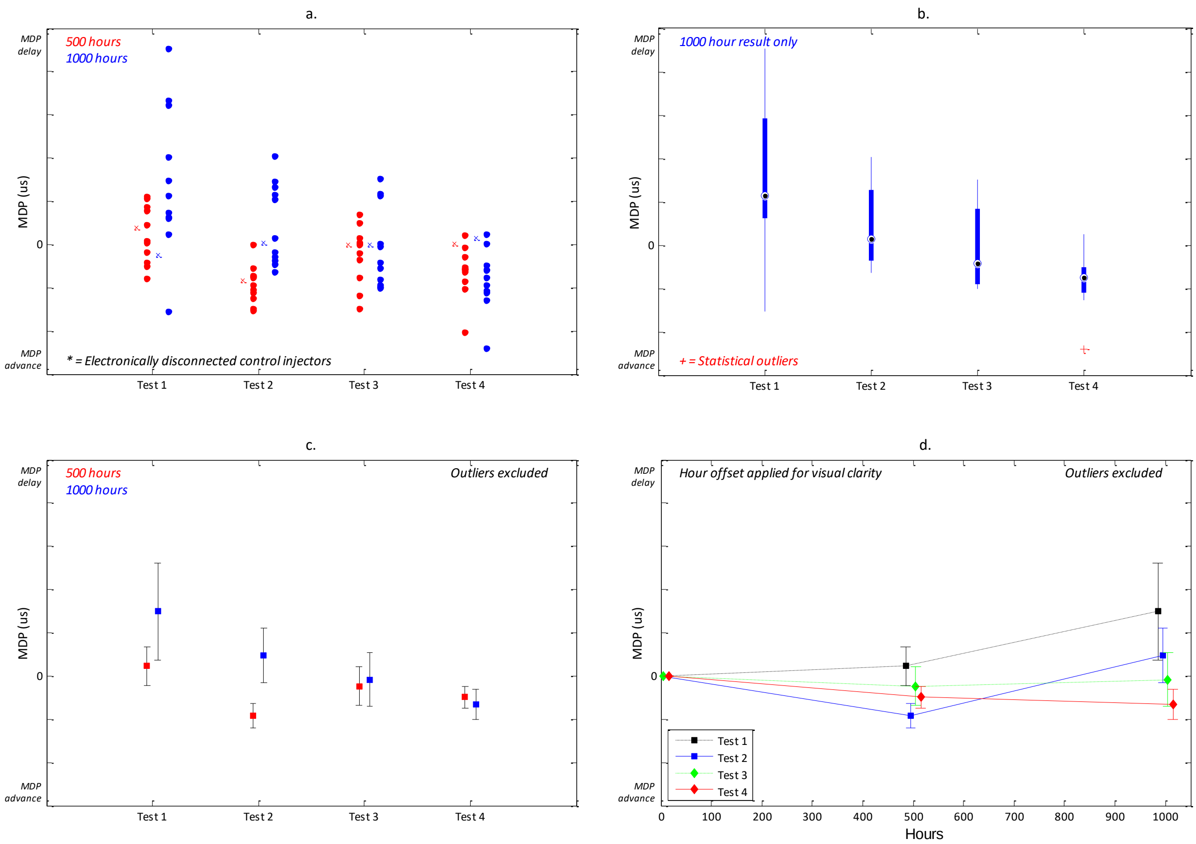


Figure 77: Summary of MDP performance metric: a.) individual values, b.) Box plot, c.) mean with 95%CI, d.) progression over time

7.7.2 Untrimmed Maximal Drift – Ballistic region

The results for the UM Drift metric in the ballistic region of the gain curve for each treatment combination are presented in Figure 78. The results for treatment combination 1 demonstrate significantly more dispersion at 1000 hours when compared to the other treatment combinations after outliers have been identified. The progression over time curves generally demonstrate a non-significant increase between 500 and 1000 hours, suggesting that when considered as UM ballistic drift, the majority of the change occurs in the first 500 hours. Treatment combination 1 demonstrates a different trend to the other combinations, showing a reversal between 500 and 1000 hours, and ultimately resulting in a negative maximal ballistic drift in the untrimmed condition. The control samples can be observed to exhibit lower change over time, but differences in the observations associated with each treatment combination can still be observed, where the control samples from tests 2 and 4 demonstrate an increased UM drift when compared to tests 1 & 3.

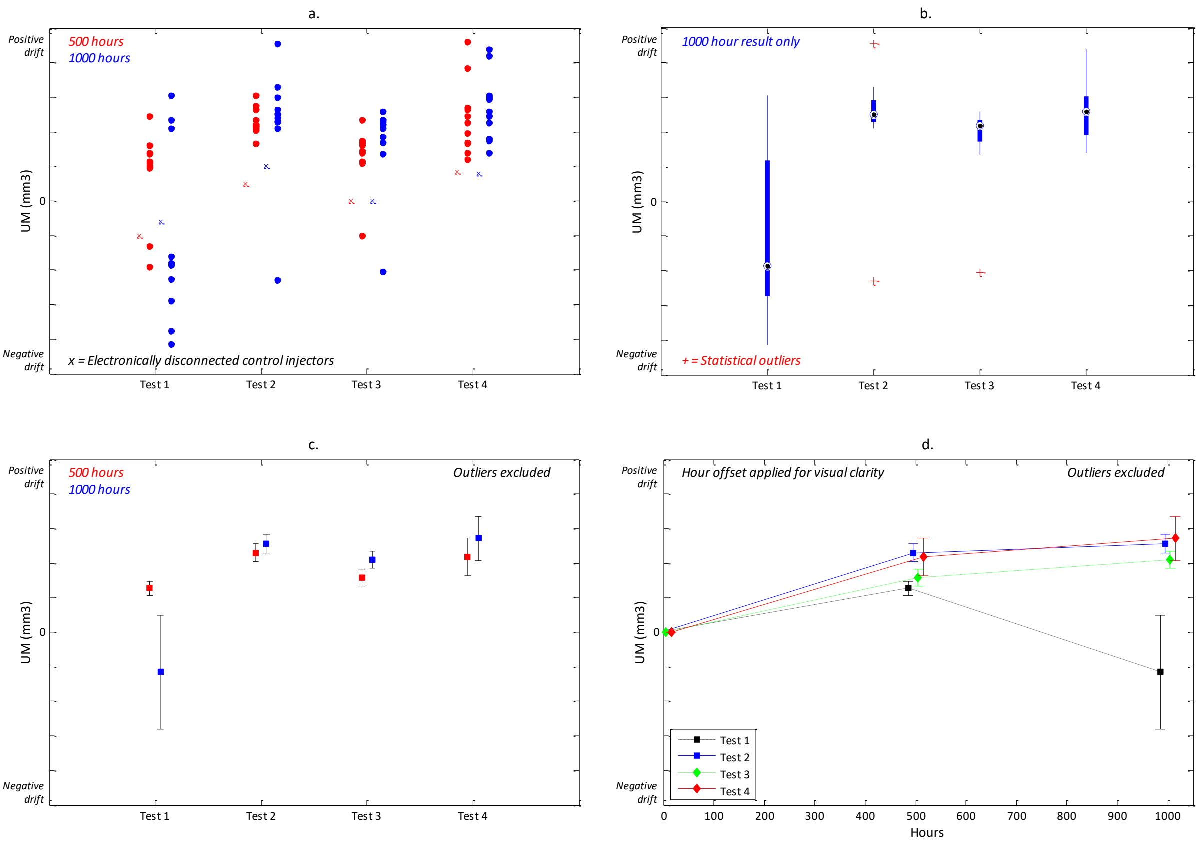


Figure 78: Summary of UM performance metric: a.) individual values, b.) Box plot, c.) mean with 95%CI, d.) progression over time

7.7.3 Trimmed Integral Drift – Ballistic Region

The results for the TI Drift metric in the ballistic region of the gain curve for each treatment combination are presented in Figure 79. Treatment combinations 2 & 4 demonstrate a significantly higher average sample mean when compared to the combinations 1 & 3. Each treatment combination exhibits equivalent levels of dispersion. While the sample mean can be seen to change for each treatment combination, only treatment combination 2 results in a significant increase in TI drift between 500 & 1000 hours. The control samples can be observed to exhibit a lower magnitude of TI drift than the test samples but exhibit a similar location trend to the corresponding sample means.

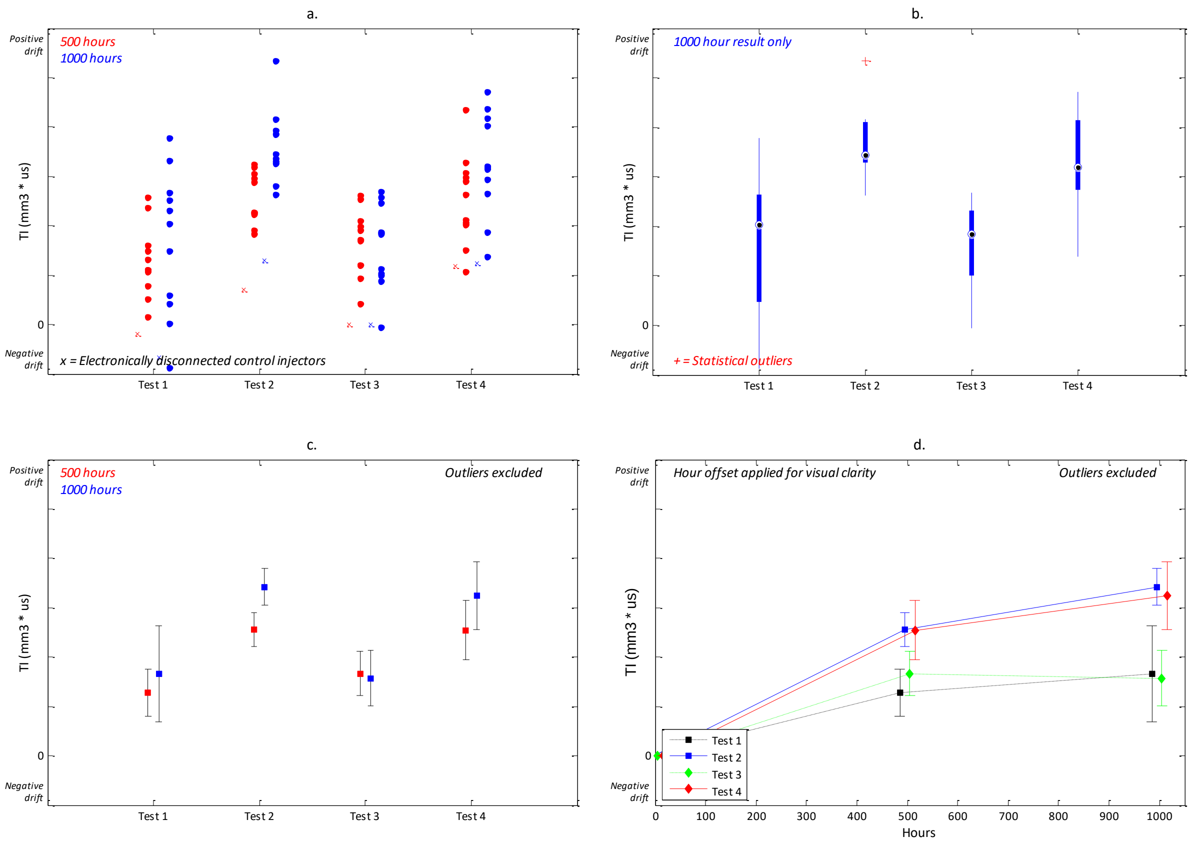


Figure 79: Summary of TI performance metric: a.) individual values, b.) Box plot, c.) mean with 95%CI, d.) progression over time

7.8 Measurements

7.8.1 NCV Seat Wear

The results for the NCV Seat Wear metric for each treatment combination are presented in Figure 80. Treatment combinations 2 & 4 demonstrate a significantly higher, with a 95% confidence level, average sample mean when compared to the combinations 1 & 3. Treatment combinations 1 & 3 exhibit a higher level of dispersion than combinations 2 & 4. The control samples can be observed to exhibit a lower magnitude of NCV seat wear than the test samples, but generally exhibit a similar location trend to the corresponding sample means.

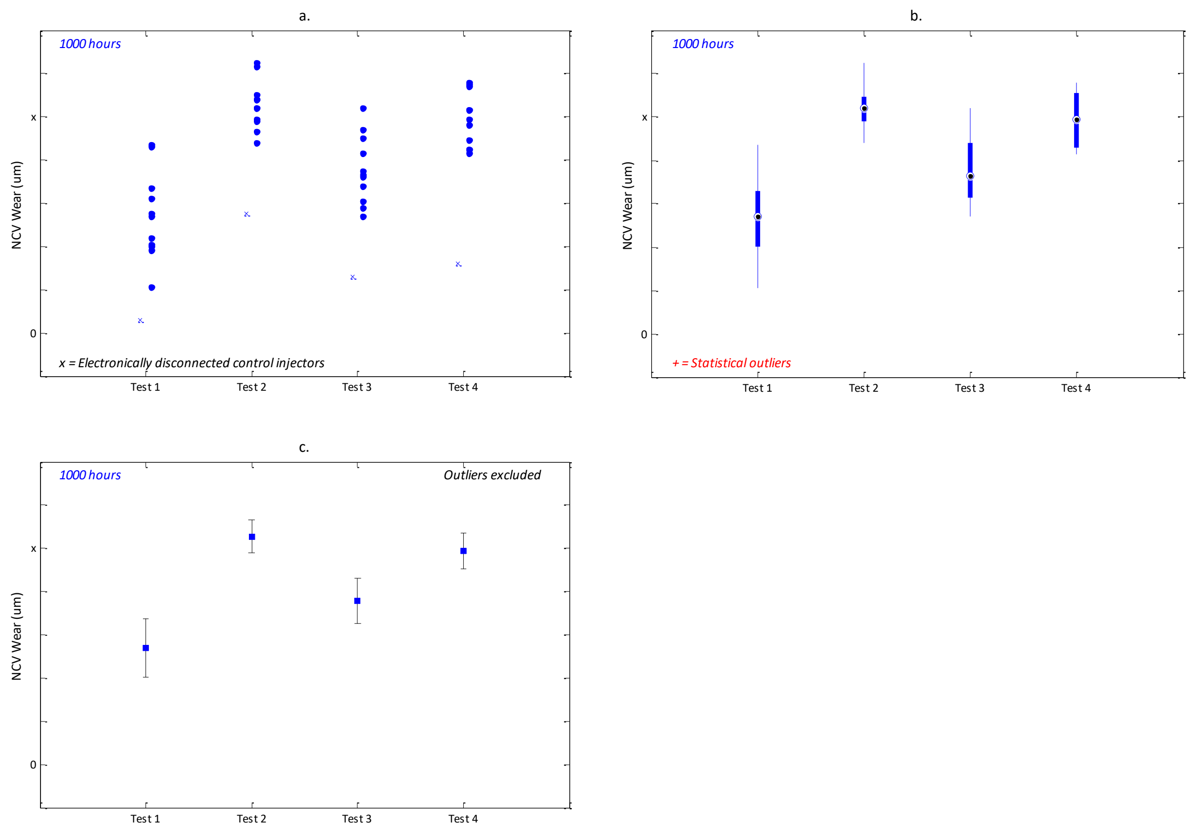


Figure 80: Summary of NCV seat wear measurement metric

7.8.2 PG trench depth

The results for the PG trench depth metric for each treatment combination are presented in Figure 81. Treatment combination 2 demonstrates a significantly higher, at a 95% confidence level, sample mean than the other combinations, while also possibly resulting in higher dispersion. The control samples can generally be observed to exhibit no material removal, with the exception of the control sample associated with treatment combination 2, which exhibits a lower magnitude of wear than the test samples, while the overall location trend of the control samples is similar to that of the test samples.

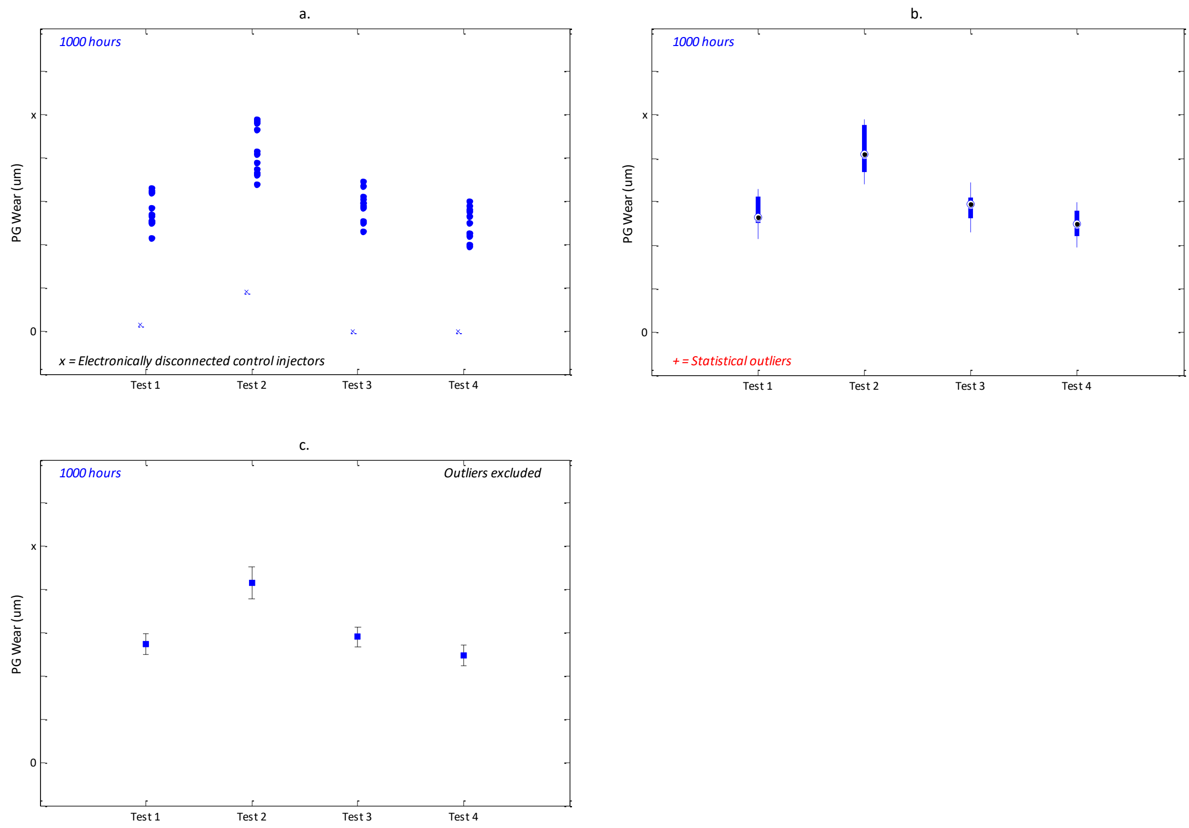


Figure 81: Summary of PG trench depth performance metric

7.9 DoE Results – Location & Dispersion Effects

7.9.1 MDP timing change

As shown in Figure 82, at the 50% confidence level, the only significant location effect observed is associated with the Fuel Type. The use of WWLTF fuel is shown to result in a positive change on MDP timing, representing a delay in start of injection. This may be indicative of a IDID effect, or another fuel related mechanism. Higher levels of both the Rail Pressure and Number of Injections variables resulted in a decrease in MDP, or an injection advance, but are not significant at the 50% confidence level.

Figure x also shows the interactions associated with each variable pairing on the MDP timing change metric. As can be seen, a strong interaction is observed between the Number of Injections, and both the Fuel type and the Rail Pressure.

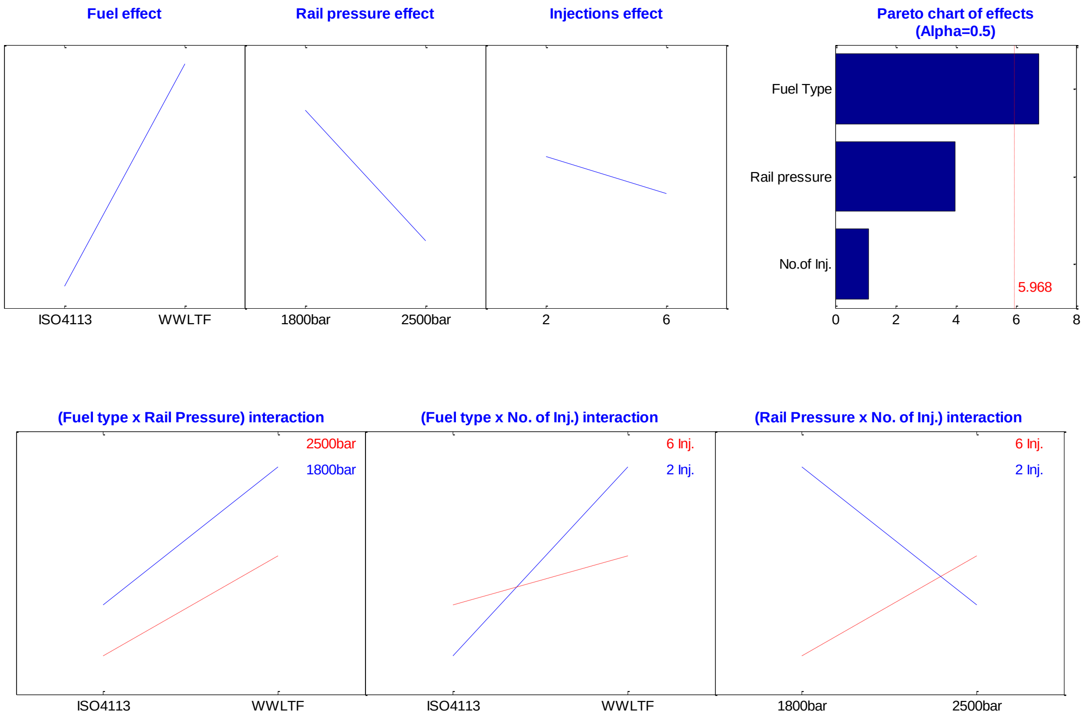


Figure 82: Visual summary of location effects and interactions for MDP performance metric

Figure 83 shows the main effects and interactions associated with dispersion for the MDP timing change metric. As can be seen, no dispersion effect is significant at the 50% confidence level, with the Fuel Type and Rail Pressure representing the two largest responses.

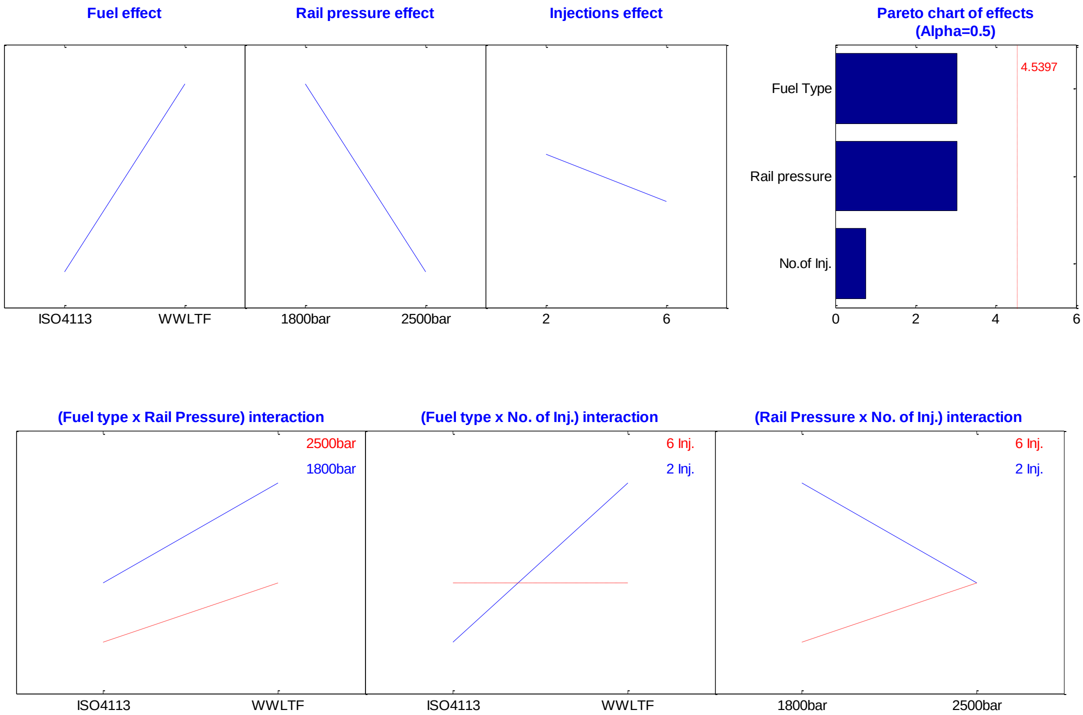


Figure 83: Visual summary of dispersion effects and interactions for MDP performance metric

7.9.2 Untrimmed Maximal Drift

Figure 84 shows the main effects and interactions associated with the location response of the Untrimmed Maximal Drift metric. As can be seen, no effects are shown to be significant at the 50% confidence level, with each effect being equivalent in relative magnitude. Increasing the level of both the Rail Pressure and Number of Injections variables may result in a non-significant increase in UM drift, while the use of WWLTF may result in a decrease in UM drift. This is contrary to the expectation that the use of lower lubricity, lower viscosity fuel would accelerate the wear of the NCV seat, perhaps suggesting either no causal relation, or a contrary wear mechanism associated with fuel degradation and IDID formation. The interaction plots suggest there is a degree of interaction associated with each variable pairing, perhaps further supporting the supposition that this metric is confounded by multiple wear mechanisms.

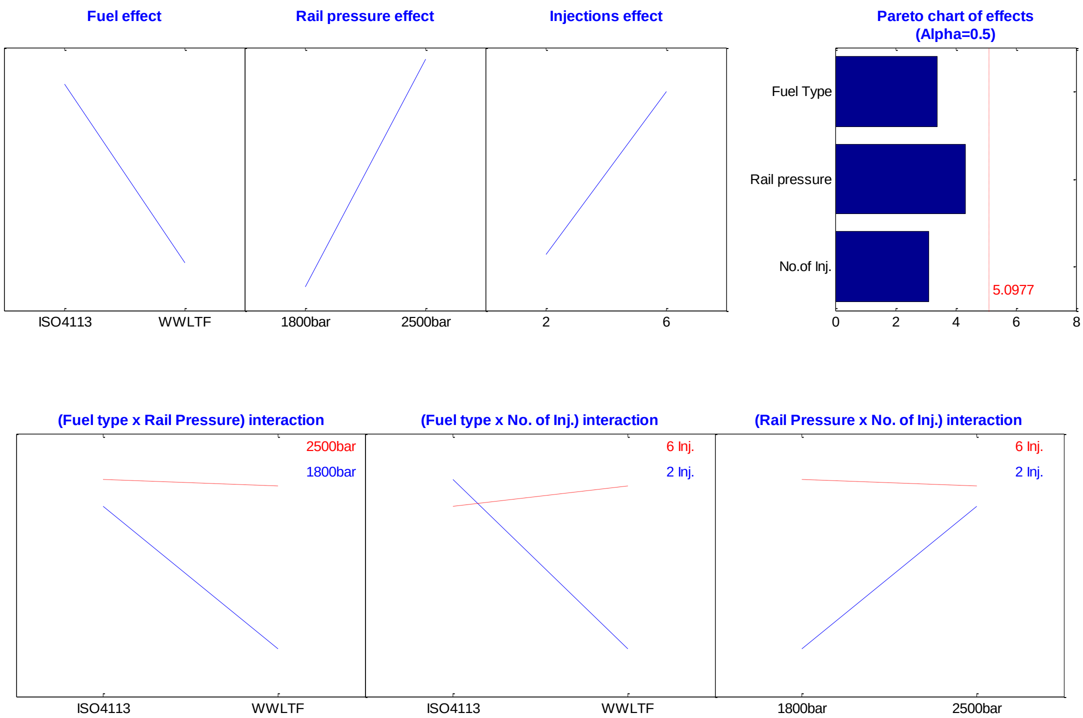


Figure 84: Visual summary of location effects and interactions for UM performance metric

Figure 85 shows the main effects and interactions associated with the dispersion response of the UM drift metric. As can be seen, the effect associated with the Number of Injections factor is significant at the 50% confidence level, with an increase in the number of injections resulting in a decrease in dispersion. A degree of interaction is observed between each factor pairing.

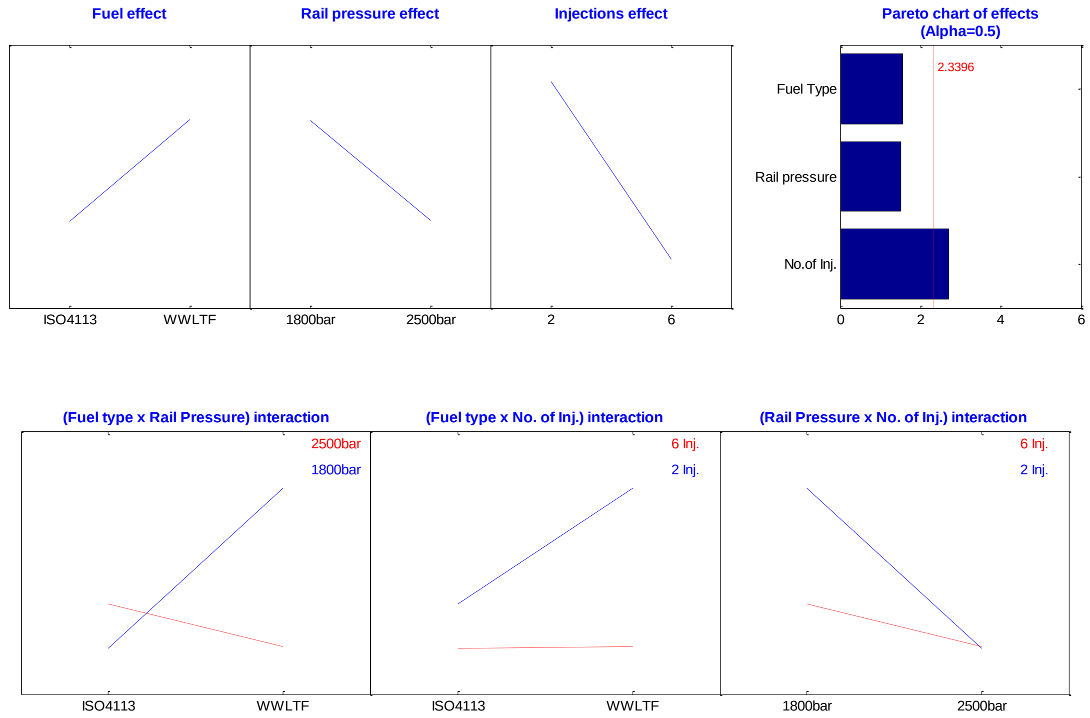


Figure 85: Visual summary of dispersion effects and interactions for UM performance metric

7.9.3 Trimmed integral drift metric

Figure 86 shows the main effects and interactions associated with the location response of the Trimmed Integral of Drift metric. As can be seen, the only significant effect at the 50% confidence level is associated with the Rail Pressure factor, and an increase in the Rail Pressure results in an increase in the TI drift metric. A strong interaction is demonstrated between the Fuel Type and Number of Injection variables.

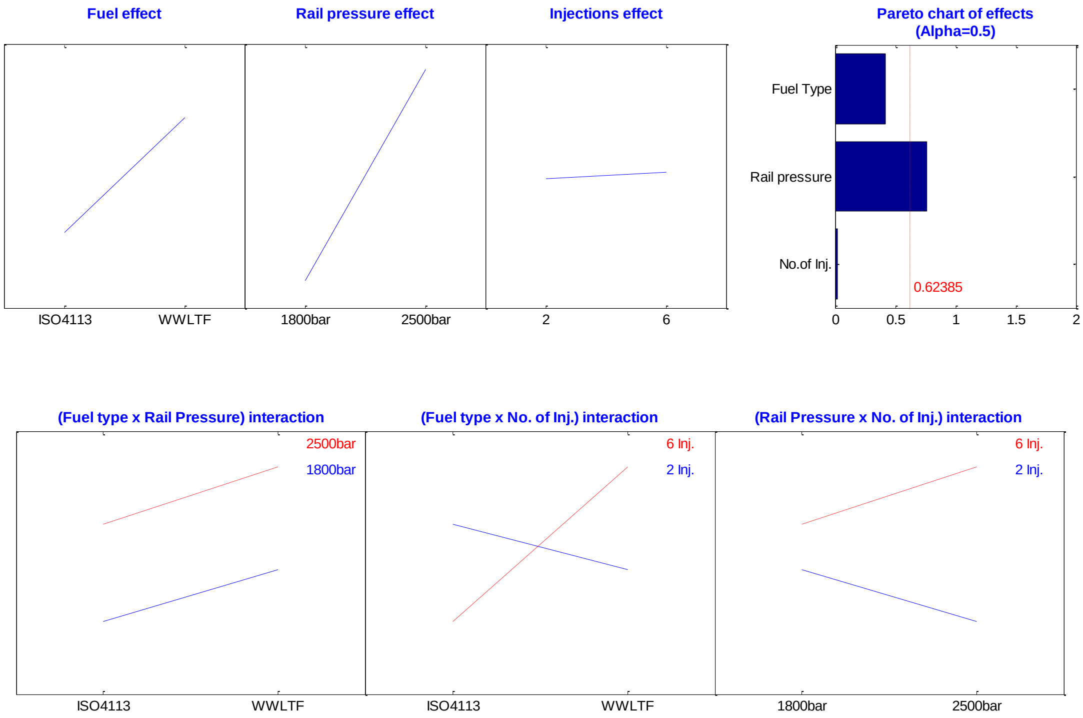


Figure 86: Visual summary of location effects and interactions for TI performance metric

Figure 87 shows the main effects and interactions associated with the dispersion response of the Trimmed Integral of Drift metric. As can be seen, the only significant effect observed is associated with the Number of Injections factor, where an increase in the number of injections results in a decreased dispersion response. Interactions can be observed to varying degrees with all variable pairs, with the strongest interaction associated with the interaction between Fuel Type and Rail Pressure.

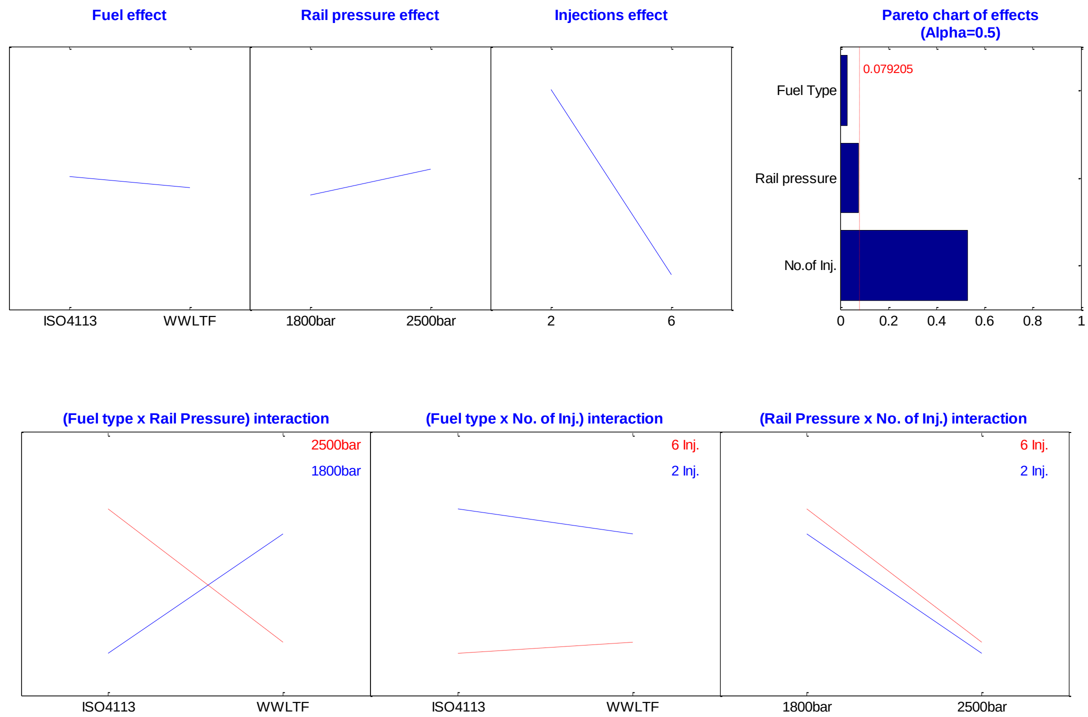


Figure 87: Visual summary of dispersion effects and interactions for TI performance metric

7.9.4 NCV Seat Wear

Figure 88 shows the main effects and interactions associated with the location response of the NCV seat wear metric. As can be seen, only the effect associated with the Rail Pressure factor is significant at the 50% confidence level and increasing the level of the rail pressure factor results in an increased level of NCV seat wear. Of the non-significant responses, Fuel Effect shows a possibly contrary response to that anticipated, where the use of more aggressive fuel may result in a decrease in wear. A strong interaction is present between the Fuel Type and Number of Injections factors, although both are insignificant.

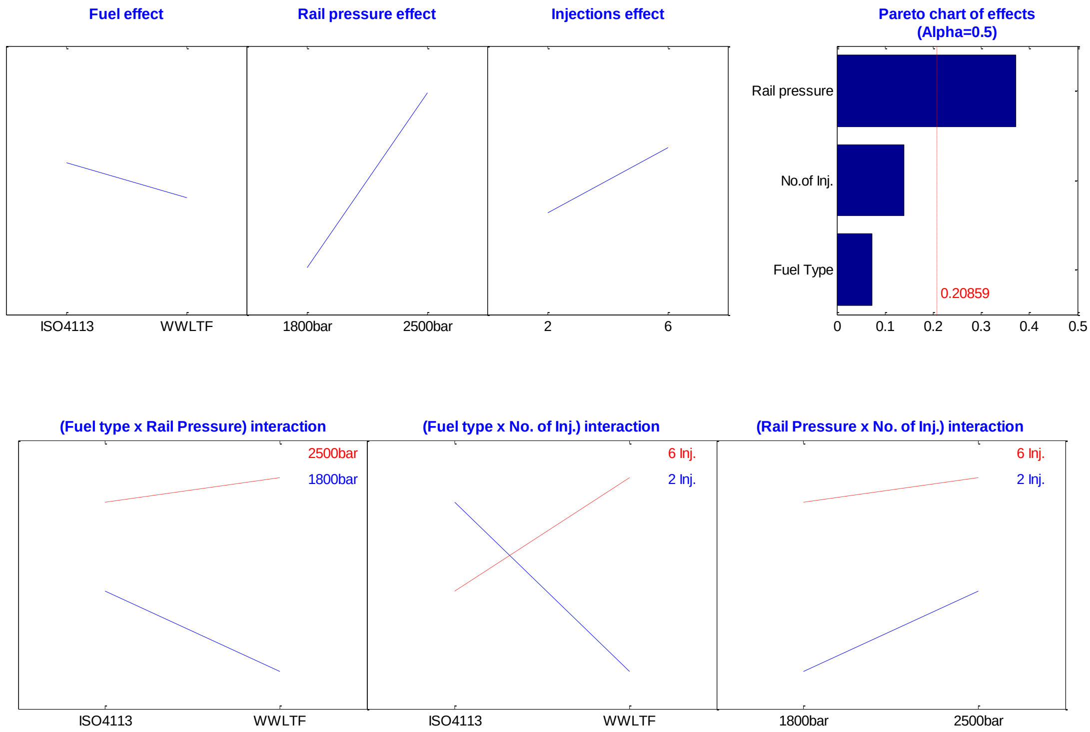


Figure 88: Visual summary of location effects and interactions for NCV seat wear measurement metric

Figure 89 shows the main effects and interactions associated with the dispersion response of the NCV seat wear metric. As can be seen, only the effect associated with the Rail Pressure factor is significant at the 50% confidence level and increasing the level of the rail pressure factor results in a decreased dispersion of NCV seat wear. A strong interaction is present between the Fuel Type and Number of Injections factors, although both are insignificant.

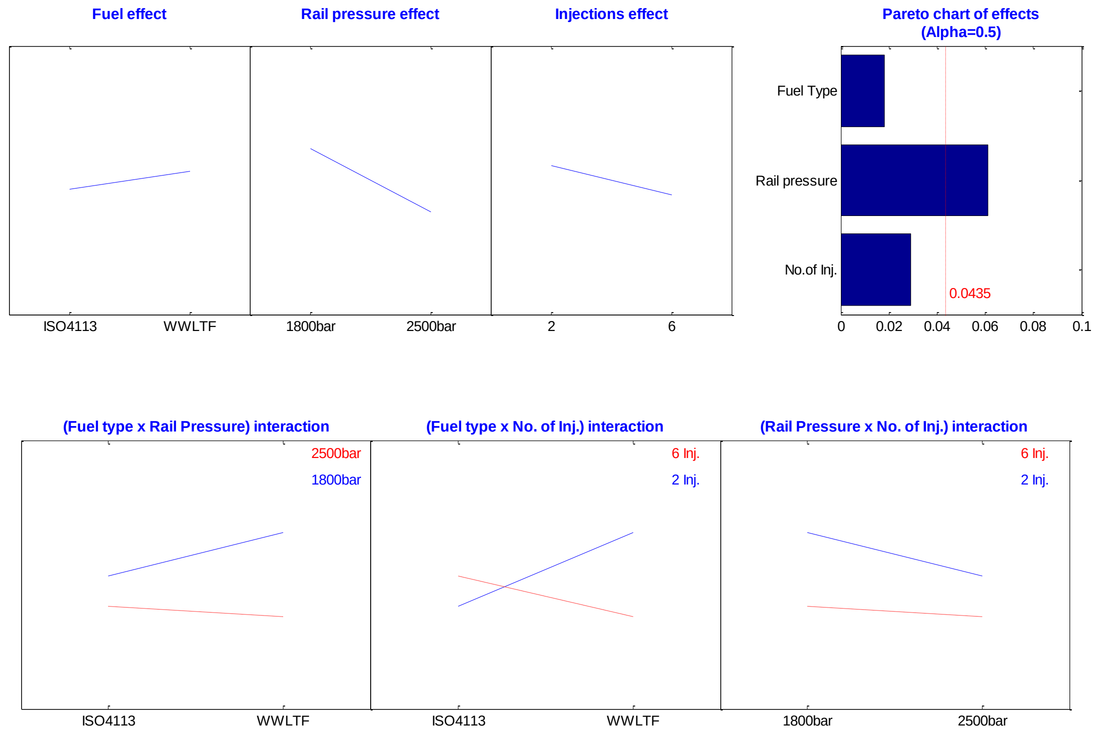


Figure 89: Visual summary of dispersion effects and interactions for NCV seat wear measurement metric

7.9.5 PG trench depth

Figure 90 shows the main effects and interactions associated with the location response of the PG trench depth. As can be seen, no effects are significant at the 50% confidence level. However, the results suggest that increasing the level of each factor may results in an increase dispersion of material removal on the PG, and the largest effect may be associated with the number of injections. A degree of interaction may also be present between each variable pairing.

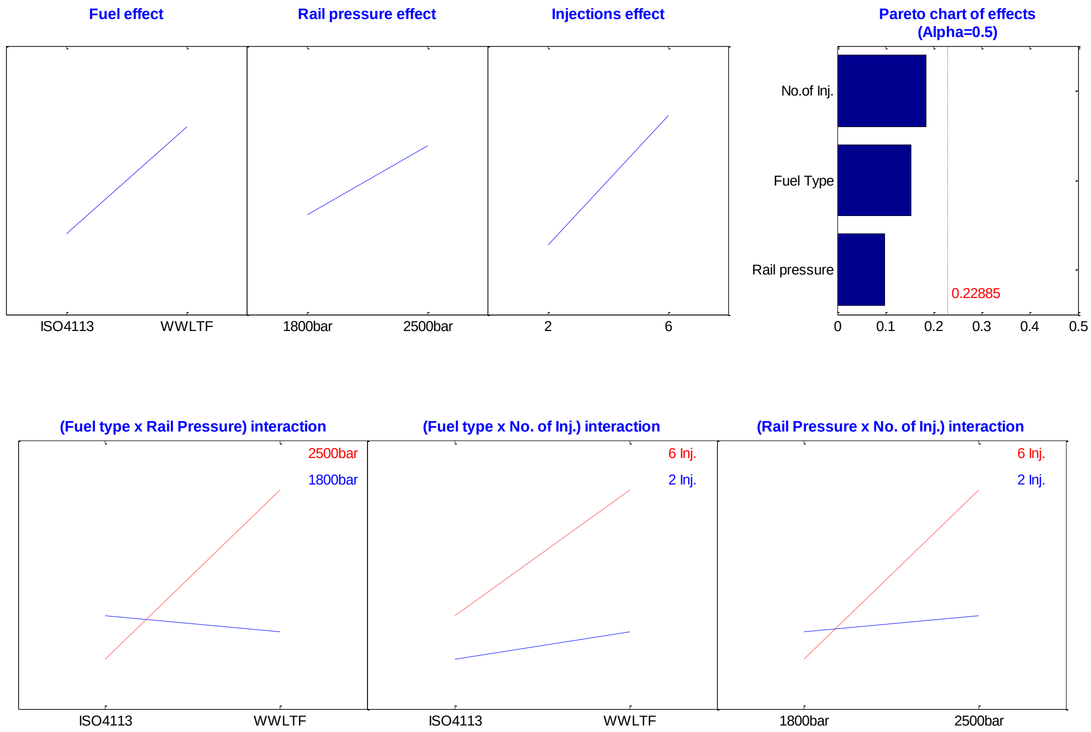


Figure 90: Visual summary of location effects and interactions PG trench depth measurement metric

Figure 91 shows the main effects and interactions associated with the dispersion response of the PG trench depth. As can be seen, no effects are significant at the 50% confidence level. However, the results suggest that increasing the level of each factor may results in an increase level of material removal on the PG, and the largest effect may be associated with the fuel. A degree of interaction may also be present between each variable pairing.

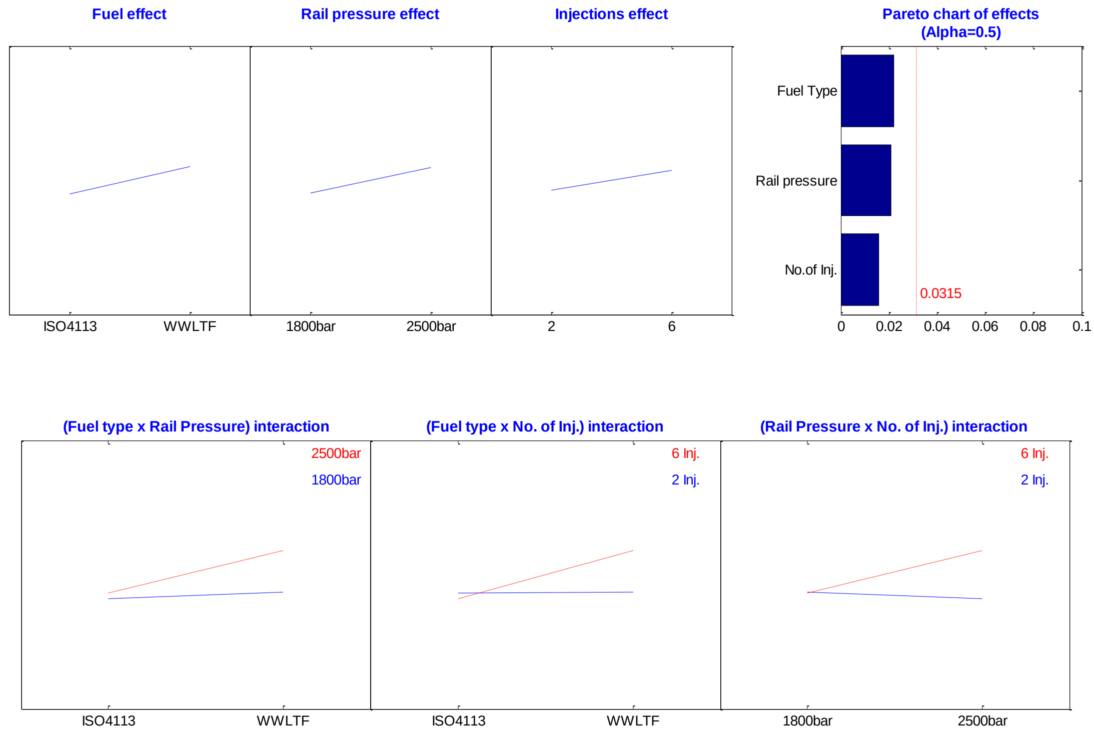


Figure 91: Visual summary of dispersion effects and interactions for PG trench depth measurement metric

7.9.6 Summary – Location effects and interactions

Figure 92 provides a visual summary of the location effects observed for each metric, allowing direct comparison of the performance metrics to those associated with material removal. As can be seen, TI Drift exhibits the same directionality and proportionality of effects as NCV seat wear, with a statistically significant effect associated with the level of Rail Pressure demonstrated for both metrics, where increasing rail pressure resulted in an increase in the observed TI Drift and NCV seat wear.

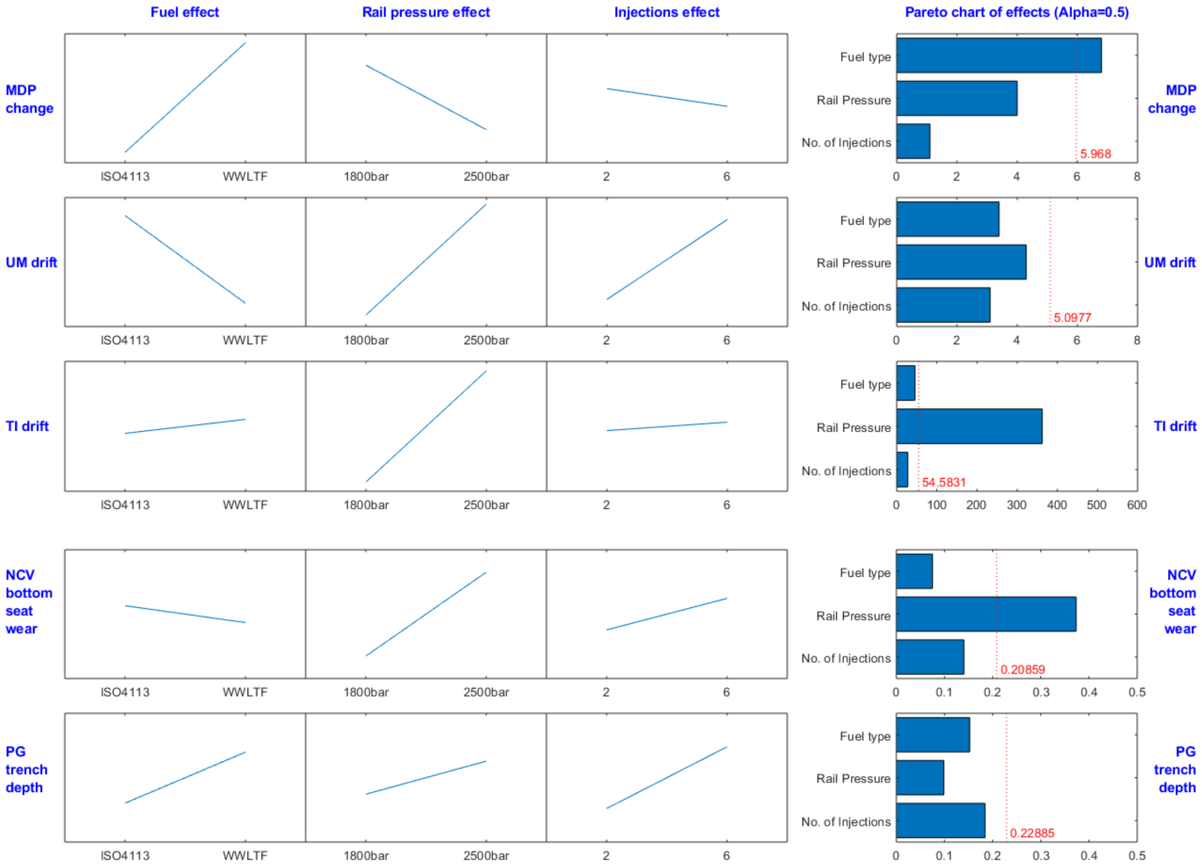


Figure 92: Visual summary of all location effects

Figure 93 provides a visual summary of the interactions between design variables observed in the location of each metric. In a similar manner to that observed in the main effects, TI drift is the only performance metric that exhibits interactions similar to those observed in NCV bottom seat wear, while MDP and Um drift exhibit alternate interactions.

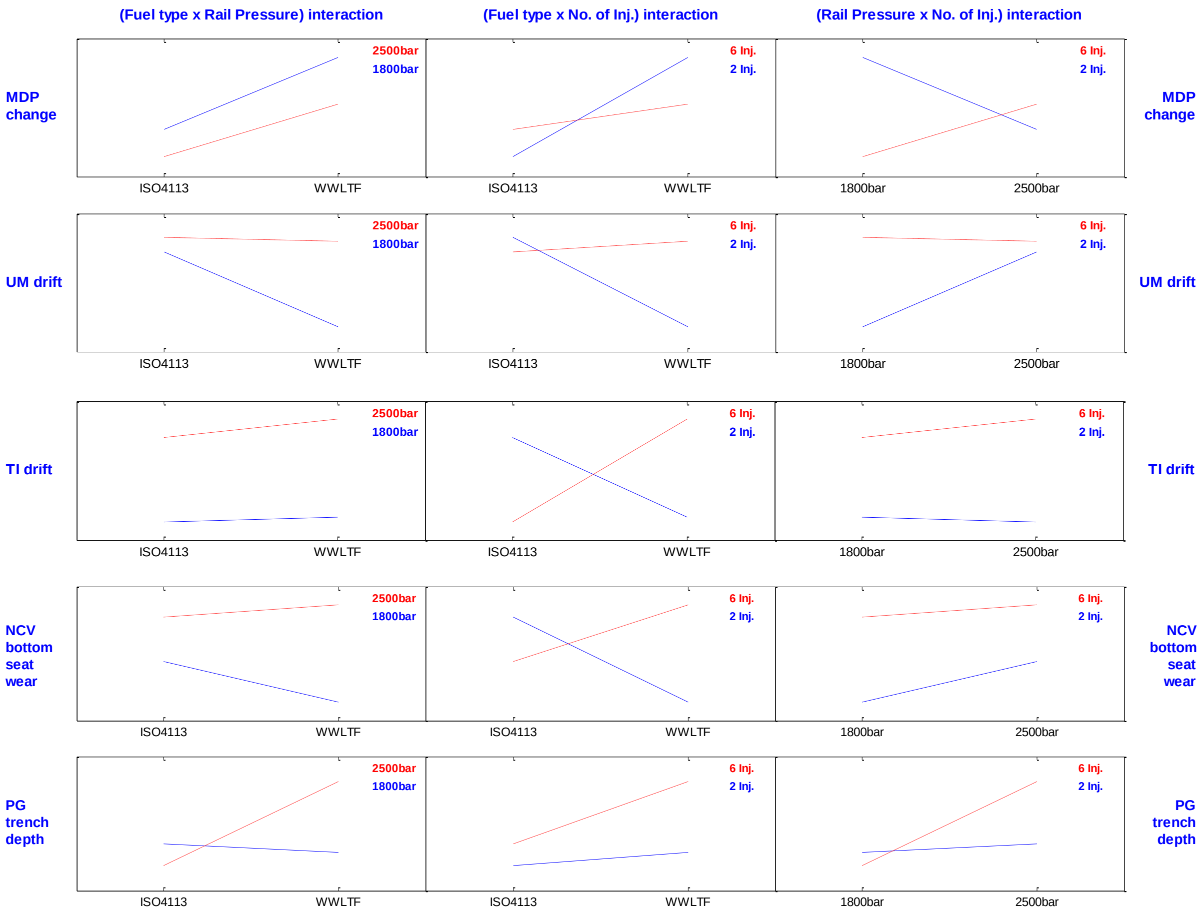


Figure 93: Visual summary of all location interactions

7.9.7 Summary – Dispersion effects and interactions

Figure 94 provides a visual summary of the dispersion effects observed for each metric, allowing direct comparison of the performance metrics to those associated with material removal. As can be seen, no performance metric corresponds entirely with the response of the NCV seat wear metric. Each performance metric exhibits the same directionality in responses as the NCV seat wear metric, but none demonstrate a significant reduction in dispersions associated with the rail pressure effect.

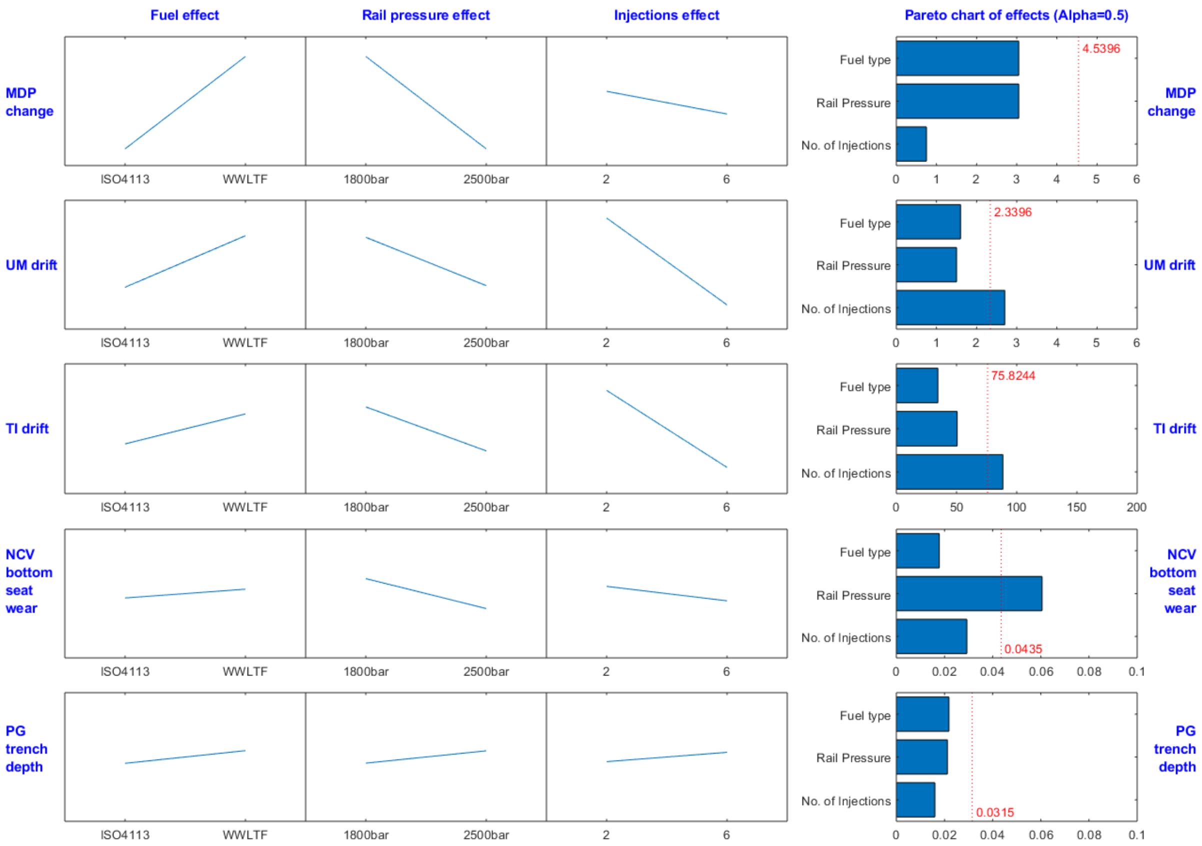


Figure 94: Visual summary of all dispersion effects

Figure 94 provides a visual summary of the interactions between design variables observed in the dispersion of each metric.

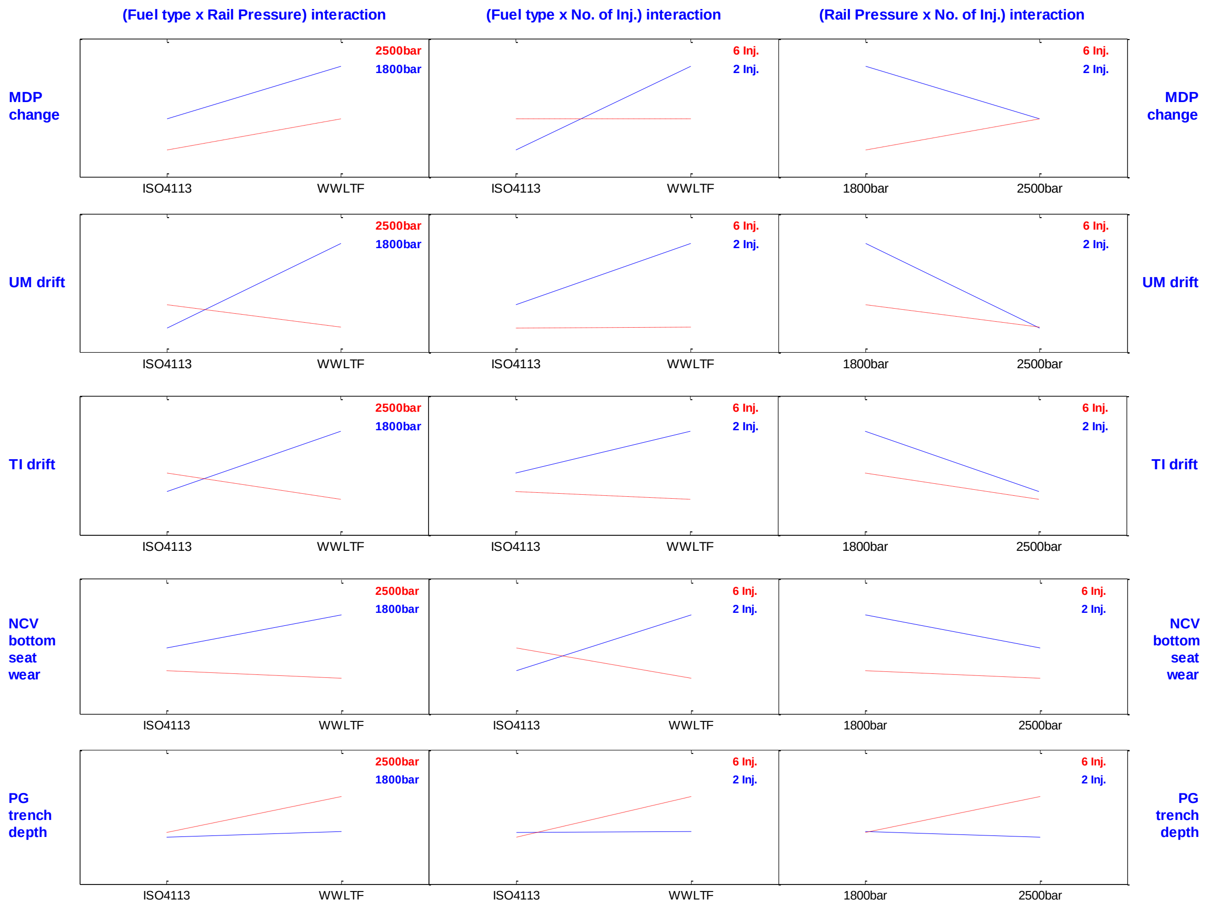


Figure 95: Visual summary of all dispersion interactions

7.10 Using design projection to determine experimental error

When considering the results of a fractional factorial experiment, the least significant design factor can be removed to provide an estimation of either the experimental error, or the significance of any interactions, using design projection (Montgomery, 2009). In the case of this 2^{3-1} fractional factorial design, removing one design factor will result in a 2^2 design, as visualised in Figure 96.

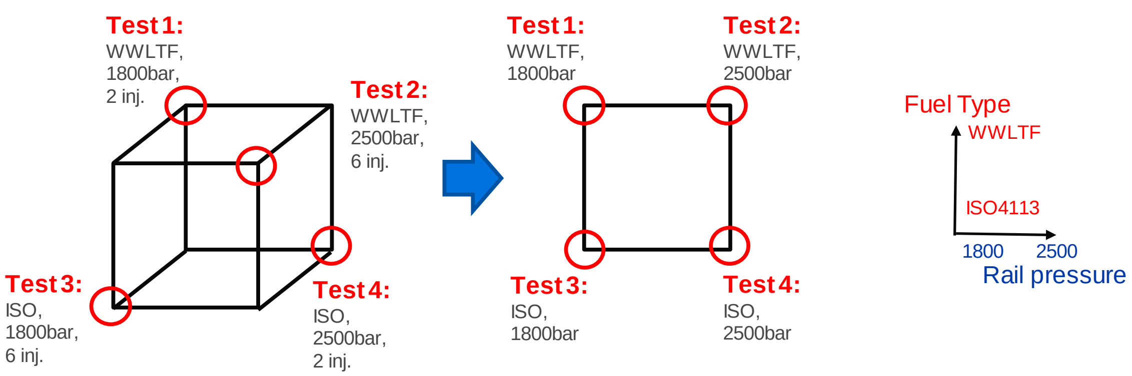


Figure 96: Visualisation of design projection from 2^{3-1} to 2^2

Considering injector drift as measured by the Trimmed Integral of Drift metric, the least significant design factor is the number of injections, removal of which would result in a 2^2 with fuel type and rail pressure as the two design variables. Analysis of this design through ANOVA is presented in Table 18. As can be seen, the Rail Pressure factor can be demonstrated to be a significant effect with >95% confidence (P-value 0.048). Similarly, the sum of squares of the Fuel Type factor effect (2078) is shown to be only ~3x greater than the experimental error (739), and not a significant effect.

Source	Degrees of Freedom	Adj. SS	F-Value	P-Value
Model	2	133302	90.17	0.074
Linear	2	133302	90.17	0.074
Fuel Type	1	2078	2.81	0.342
Rail Pressure	1	131224	177.52	0.048
Error	1	739		
Total	3	134041		

Table 18: ANOVA for projected 2^2 design

Using the projected 2^2 design, Equation 2 can be used to describe for TI drift can be describes in terms of fuel type and rail pressure.

$$TI = -608.2 + 22.8 \times Fuel\ Type + 0.5175 \times Rail\ Pressure \tag{2}$$

This equation can be shown to have a high degree of correlation to the experimental results as shown in Table 19.

S	R^2	R^2(Adj).	R^2(Pred.)
27.1881	99.45%	98.35%	91.18%

Table 19: Correlation of projected 2^2 design to experimental results

7.11 Test 5 – Linearity of Rail Pressure Effect on Injector Performance

With rail pressure being identified as the most significant effect for both TI drift and NCV seat wear, there is significant interest in exploring the linearity of its effect. The testing facilities, and unused test samples, remained available to support limited additional testing, so the decision was made to run a 5th test incorporating a centre point for rail pressure. Test 5 would be identical in sample allocation, total duration, and measurement to the first 4 tests, but would feature an additional performance characterisation interval at a low hour (<100), as will be discussed in Chapter 9. Table 20 then represents an update of the experimental structure described previously in Table 16 with the addition of Test 5.

Test Feature	Count, duration, or time interval
Number of tests	5
Number of samples per test *(including 1x electronically disconnected sample per test)*	12
Duration of each test	1000 Hours
Performance characterisation intervals for each sample and test *(*applicable for Test 5 only)*	0 Hours 67 Hours (*) 500 Hours 1000 Hours
Disassembly and component measurement interval for each sample and test	1000 Hours

Table 20: Test count, sample allocation, duration, and measurement intervals with the addition of Test 5

Using the projected 2^2 factorial design presented in §7.10, Fuel Type represents a discrete design factor, and as such, 2 pseudo centre points would be required for a fully orthogonal design, with each fuel tested at 2150bar. This is visualised in Figure 97.

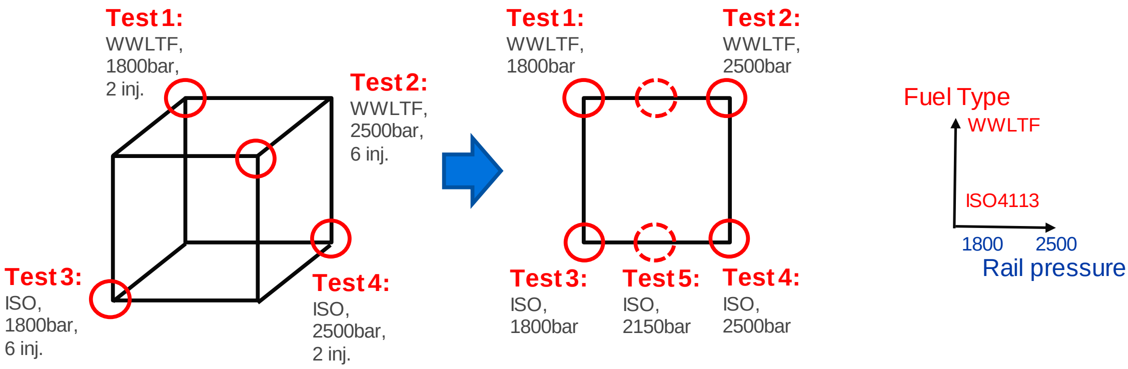


Figure 97: Visualisation of addition of pseudo centre points to projected 2^2 design

The first test to be completed used the combination of ISO4113 with 2150bar rail pressure. Unfortunately, after this time the testing facility was required to support prototype development activities, so the second pseudo centre point associated with WWLTF fuel could not be completed. However, as shown in §7.9, the fuel type effect has so far been shown to be statistically insignificant.

Using the regression equation for the projected design presented in §7.10, if a linear effect was associated with the rail pressure effect, then the expected sample mean of TI drift for this pseudo centre point after 1000 hours with ISO4113 (Fuel type = 0) and 2150bar would be found in Equation 3.

$$TI = -608.2 + 22.8 \times (0) + 0.5175 \times (2150) = 504 \tag{3}$$

Figure 98 shows the summary for each treatment combination for the TI drift metric with test 5 included. As can be seen, the sample mean for test 5 after 1000 hours is over 600 (628), representing a higher level than that predicted for a linear effect of rail pressure.

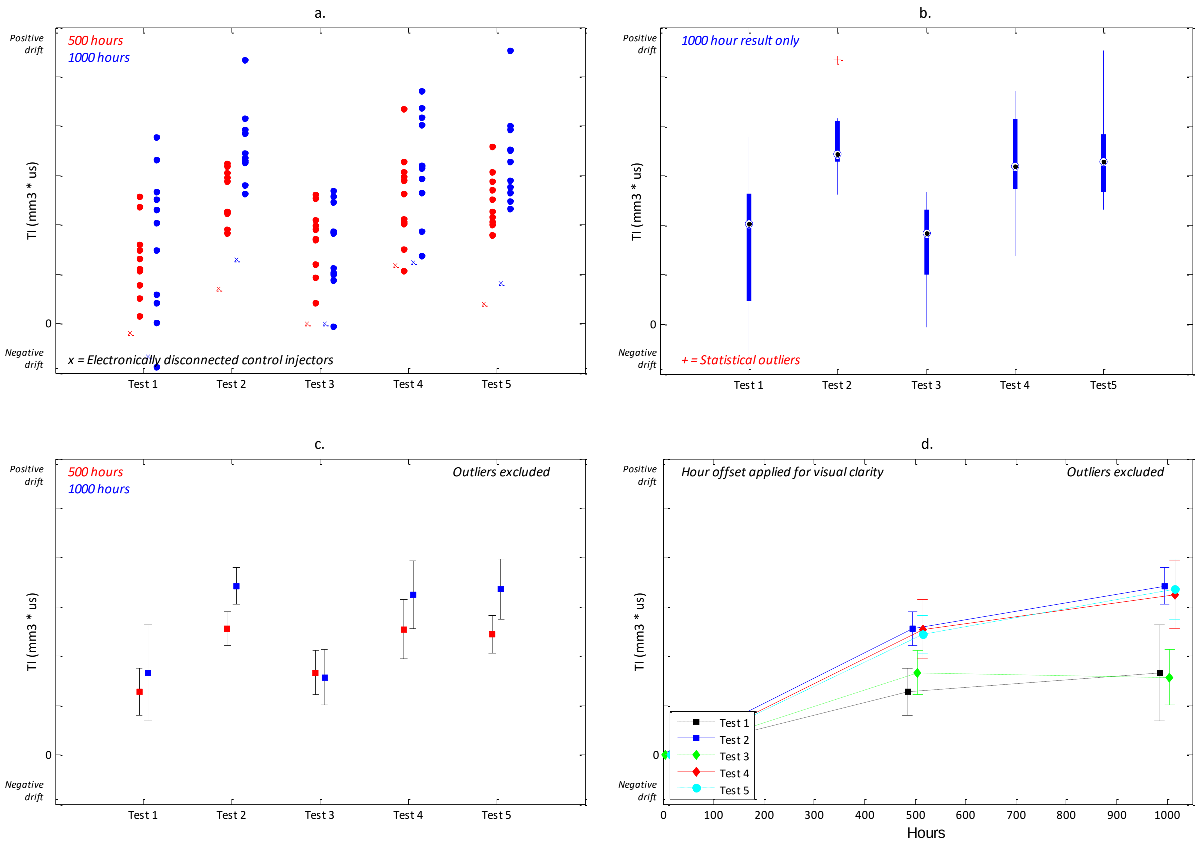


Figure 98: Summary of TI performance metric with the addition of pseudo centre point (Test 5) : a.) individual values, b.) Box plot, c.) mean with 95%CI, d.) progression over time

Analysis of the location effects is then presented in Table 21 and Figure 99. With only one centre point, the design is not orthogonal, but additional insight to the linearity of the Rail Pressure effect for TI drift can be gained.

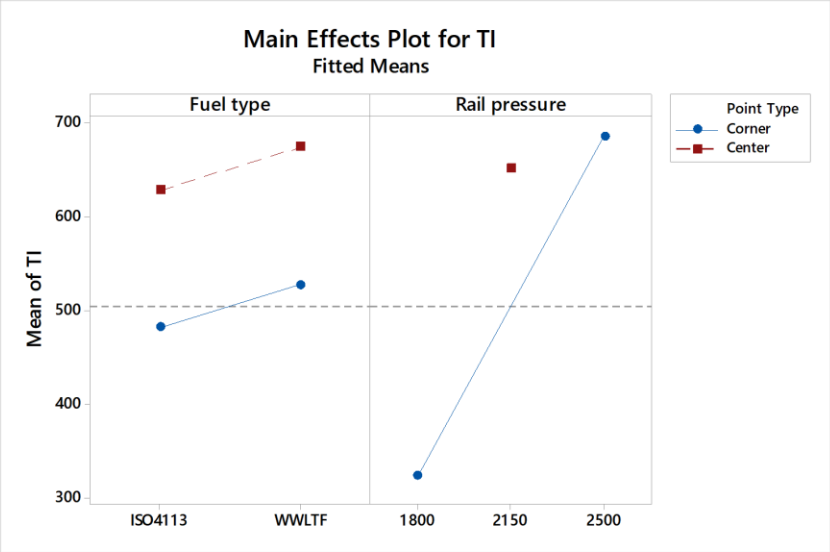


Figure 99: Visualisation of location effects with the addition of a pseudo centre point

Source	Degrees of Freedom	Adj. SS	F-Value	P-Value
Model	3	145522	65.62	0.090
Linear	2	133302	90.17	0.074
Fuel Type	1	2078	2.81	0.342
Rail Pressure	1	131224	177.52	0.048
Curvature	1	14286	19.33	0.142
Error	1	739		
Total	3	134041		

Table 21: ANOVA for the projected 2^2 design with a pseudo centre point

While the single pseudo centre point does not result in full confidence, the analysis does suggest, with a relatively high confidence of >85% (P-value 0.142), that the relationship between Rail Pressure and TI Drift is non-linear.

If the relationship between Rail Pressure and TI drift is non-linear, then a suitable regression model needs to be determined to describe that relationship as in Equation 4.

$$TI\ Drift = F(Pressure) \tag{4}$$

Table 22 presents a summary of suitable regression models, both linear and non-linear, while Figure 100 then presents a visualisation of the regression models with both the original 2^{3-1} factorial design points, and the subsequent centre point, included for reference. A linear regression model fitted to the original 4 points from the 2^{3-1} factorial design is included for reference, with the standard error for each model quoted being inclusive of the centre point.

Model	**Equation** (Drift (TI) as a function of Pressure (RP))	**Standard error**
Linear #	$TI = -608.2 + 0.5175 * (RP)$	77.65
Linear	$TI = -583.5 + 0.5175 * (RP)$	70.80
Quadratic	$TI = -5148 + 4.856 \times (RP) - 0.001009 \times (RP)^2$	37.54
Michaelis-Menten	$TI = 314 + 454.799 * \left(\frac{(RP - 1800)}{157.074 + (RP - 1800)}\right)$	31.57
Logarithmic	$TI = 314 + 82.9997 \times \ln(RP - 1800) - 172.289$	*
Power	$TI = 314 + 75.7141 * (RP - 1800)^{0.242776}$	*

 # Fitted to original 4 DoE points only but standard error calculated with addition of centre-point

** No degrees of freedom available to calculate standard error*

Table 22: Comparison of error associated different regression models

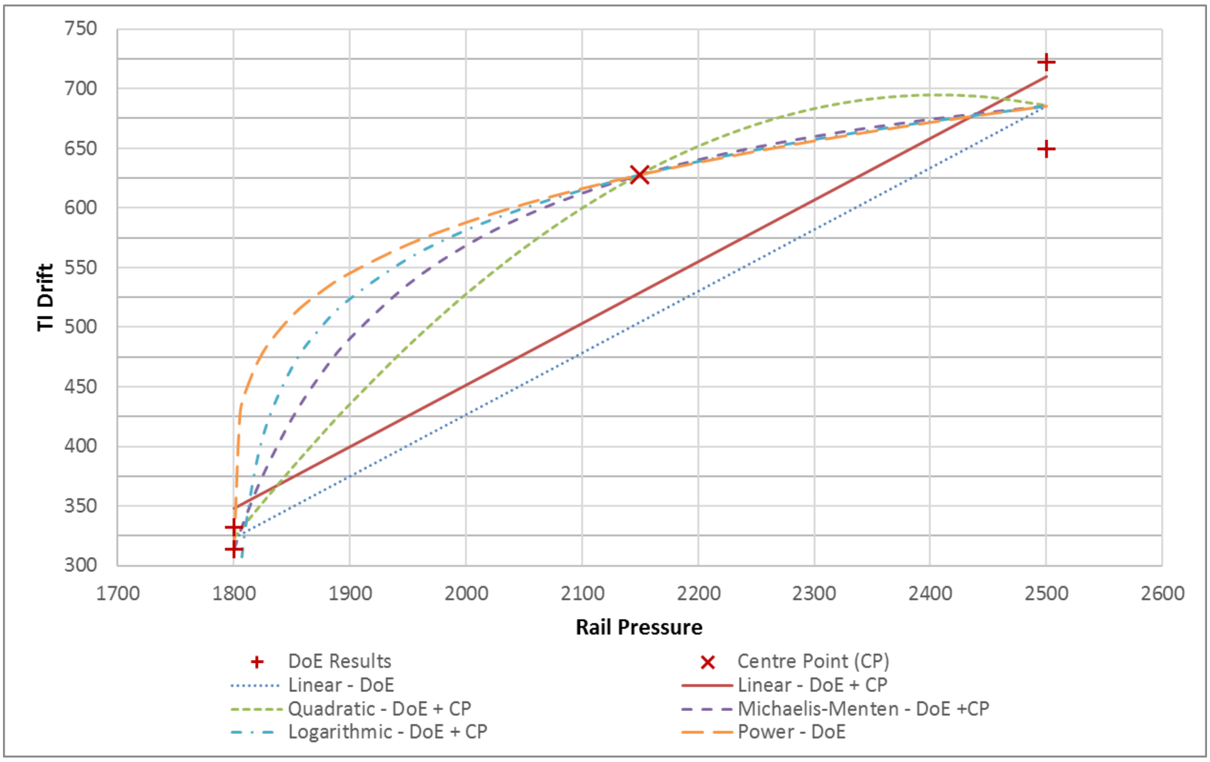


Figure 100: Visualisation of fitted regression models and experimental sample means

In the case of the Michaelis-Menten, Logarithmic, and Power regression models, an offset was required for both the values of TI and Rail pressure such that the model would pass through the origin; this offset is reflected in their equations. As can be seen, the lowest standard error is associated with the Michaelis-Menten model, an asymptotic non-linear regression model. For two of the models, Standard Error could not be calculated as no degrees of freedom remained for the calculation; additional test results could be used to test the error associated with those models, while refining the calculation for all models.

Of the non-linear regression models, the quadratic model is the only non-asymptotic model, for which values of rail pressure greater than 1800bar would result in decreasing levels of TI drift. Of the remaining non-linear regression models, the relative difference between each model can be seen to be pertaining to the shape of the curve between 1800bar and 2150bar, and the consequential fit of the model at 1800bar. While no standard error could be determined for the Logarithmic or Power regression models, they can both be seen to result in adequate fit, with different curve shapes predicted at the lower end of the prediction range. The Michaelis-Menten model, which generated the lowest standard error, can be seen to result in a more progressive curve shape between 1800bar and 2150bar.

Within the bounds of this experiment, and with a confidence of >85%, TI drift can be seen to be non-linear with increasing Rail Pressure, and the Michaelis-Menten model has been fitted to describe that relationship. Additional testing, both within, and outside of, the existing experimental space will provide additional confidence for interpolation and extrapolation. Equation 5 describes TI drift in terms of rail pressure (RP).

$$TI = 314 + 454.799 * \left(\frac{(RP - 1800)}{157.074 + (RP - 1800)}\right) \tag{5}$$

7.12 Relationship between NCV seat wear and Trimmed integral of drift

The DoE programme has identified that TI drift is the most suitable performance-based inference of NCV seat wear, both in terms of differentiating between sample populations, and in terms of significant and directionality of effects. As such, there is a clear interest in modelling the relationship between TI Drift and NCV seat wear.

Figure 101 visualises the test results from this empirical investigation, with NCV seat wear measurements plotted against TI drift. For indicative purposes, both a linear and non-linear regression model have also been plotted, but the high dispersion associated with the performance metrics result in the possibility of over-fitting the data. As such, while a non-linear regression model may be fitted to the sample population with a better statistical correlation, a linear regression model passing through the origin is presented as the most suitable approximation at this time, as described by Equation 6.

$$NCV\ seat\ wear = 0.001121 \times TI \tag{6}$$

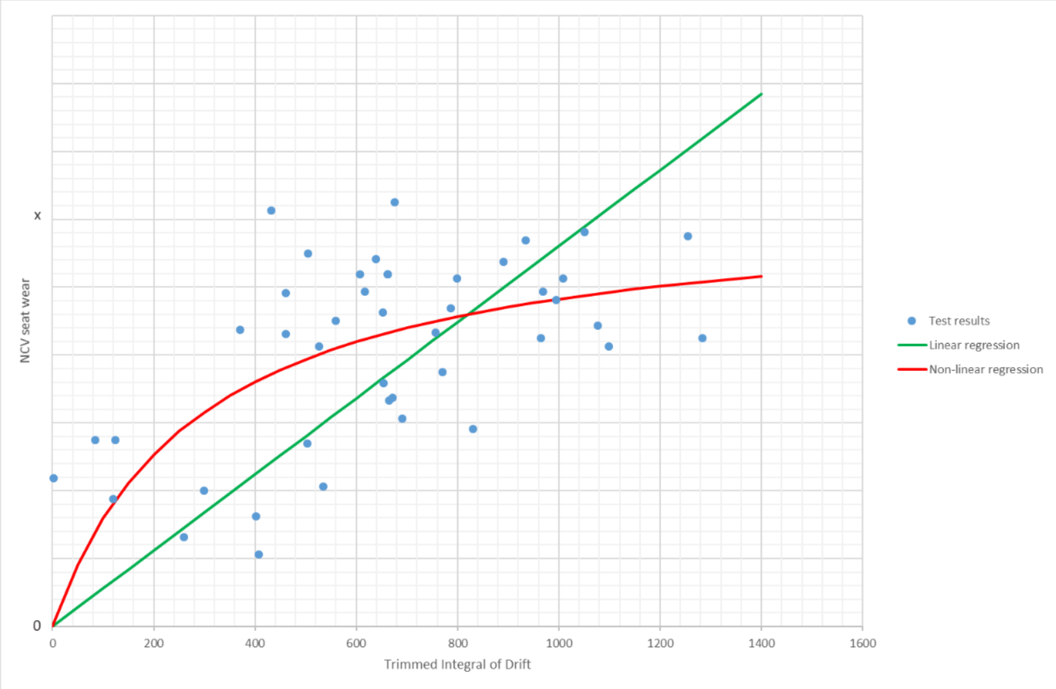


Figure 101: Visualisation of the relationship between TI drift and NCV seat wear

7.13 Summary

This Chapter has provided an overview of the second element of the FMC case study associated with this thesis. This Chapter has used the expert judgement codified in Chapter 6 as the inputs for an empirical investigation using Experimental Design methodology. The considerations associated with the design factor selection, selection and development of metrics, and experimental design have been discussed, commenting on the balance of rigour and practicality in an industrial context. A 2^{3-1} Fraction Factorial design was presented, where 4 groups of 12 samples were each tested for 1000 hours.

The results of the Design of Experiments were presented for each metric, while main effects and interactions associated with each design factor were discussed with respect to both location and dispersion effects for each metric. A single injector performance metric, trimmed integral of drift, was identified as providing a suitable inference for NCV seat wear, and the relationship between the two metrics is discussed. Finally, the 2^{3-1} design was projected to a 2^2 in order to quantify experimental error, and it was shown that only the rail pressure design factor had a statistically significant location effect on the response metrics. A pseudo centre point was then included to examine the response to Rail Pressure design factor, suggesting that the relationship was non-linear.

The next Chapter utilises the results of this Chapter, in terms of a numerical understanding of the main effect and its associate response, as an input into a regression modelling process in order to further characterise the failure mode.

Chapter 8 NCV Seat Wear – Modelling the Failure Mode

8.1 Introduction

This Chapter introduces the third and final element of the case study of the Failure Mode Characterisation method outlined in this thesis. In this element, the results of the structured empirical investigation in Chapter 7 are used as the inputs in a regression modelling process. The injector performance change over time data associated with each treatment combination is used to fit a series of non-linear regression models, with additional performance characterisation intervals used at low hours on further tests to better characterise the failure mode progression.

A generalised model of the failure mode is then developed considering those regression models, in terms of time, and the main effect identified in Chapter 7, using the Michaelis-Menten model as a suitable approximation of the failure mode. That generalised model is then validated using alternative pre-existing empirical results, in the process of which generating new knowledge about the relationship between steady state and cyclic test conditions. The generalised model is the used to compare different test cycles to application duty cycles.

The results of this element of the FMC method represent a deeper understanding of the failure mode progression, and the relationship between usage conditions and FIS life. More generally, this element of the case study presents the use of regression modelling and additional structured testing as a means for further characterising a failure mode, allowing validating of new knowledge, while providing the basis for an analytical model of the problem. An overview of this Chapter, and element of the FMC method, is visualised in Figure 102.

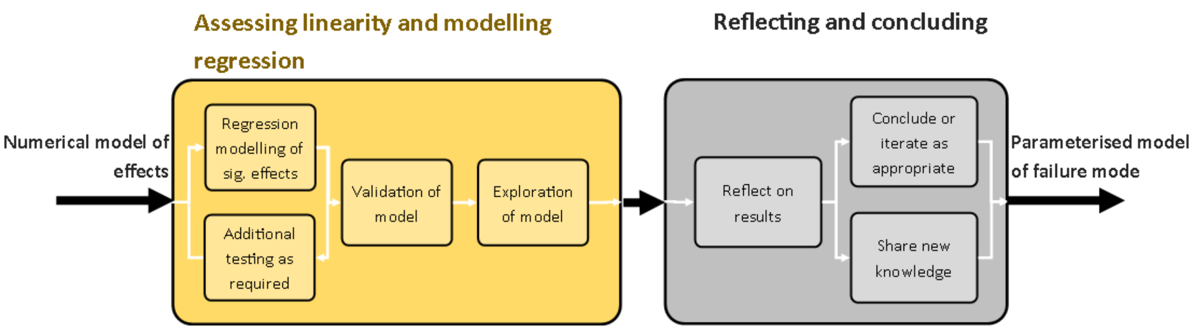


Figure 102: Overview of this Chapter

8.2 Shape of Drift over time regression curve

The 2^{3-1} factorial design yielded insight into the drift over time curve for different treatment combinations. In the case of the TI drift metric, the results were as visualised in Figure 103. For each treatment combination, a possible non-linear relationship can be observed, with the difference between 0 hours and 500 hours being greater in magnitude that the difference then between 500 and 1000 hours. This result, when viewed in conjunction with the existing expert judgement in the form of the working mental model previously presented in Figure 35, suggests a possible asymptotic relationship between Drift and hours run.

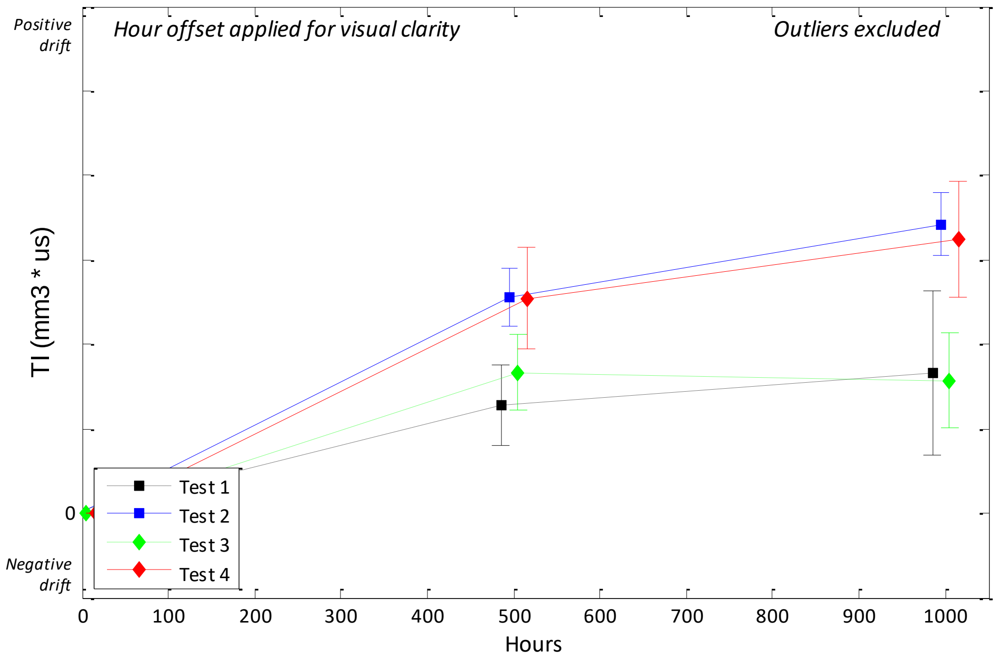


Figure 103: Progression of TI drift over time

In addition to representing a pseudo centre point, the decision was to include an additional low hour (<100) performance inspection into Test 5. Whilst it was not practical to include such an inspection into each of the original tests, the results to date suggested that valuable new information could be gained on the relationship between drift and hours run. As such, Test 5 included an additional measurement of each performance metric corresponding to 58 hours, an arbitrary time dictated by practicality rather than by design. Figure 104 includes this additional result into the drift over time curve for test 5. As can be seen, the additional inspection performed at 58 hours provides additional evidence to support the non-linearity of TI drift with respect to time.

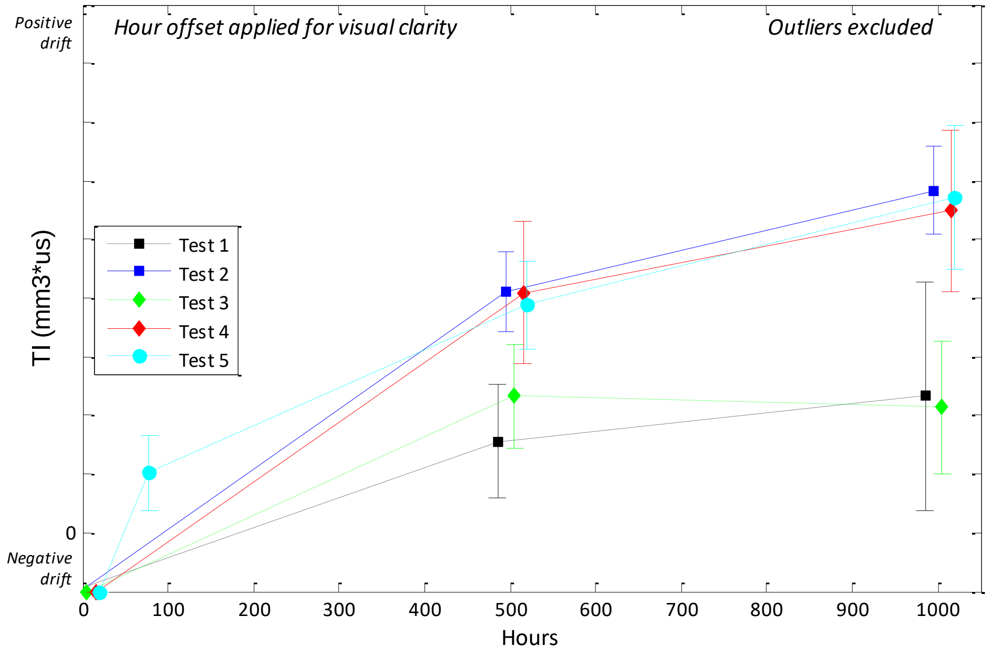


Figure 104: Progression of TI drift over time, with test 5 included

8.2.1 Fitting a regression model to the empirical data

Using the performance metrics available at the different inspection points, regression models may be developed to describe drift with respect to hours run. Firstly, considering test 2 where the highest ultimate levels of TI drift were observed, a study was completed using different regression models, observing both their statistical correlation, and their behaviour with regards to existing empirical data and expert judgement alike.

Figure 105 visualises the results of linear regression model fitting, using both a linear, and quadratic model, while Table 23 then presents the corresponding regression equations and their statistical correlations to the test data. The quadratic model is demonstrated to result in a statistically better fit to the data, providing confidence in the hypothesis that the relationship between drift and time is non-linear. However, the quadratic model would likely not be valid for extrapolation beyond 1000 hours as the model has decreasing limit behaviour.

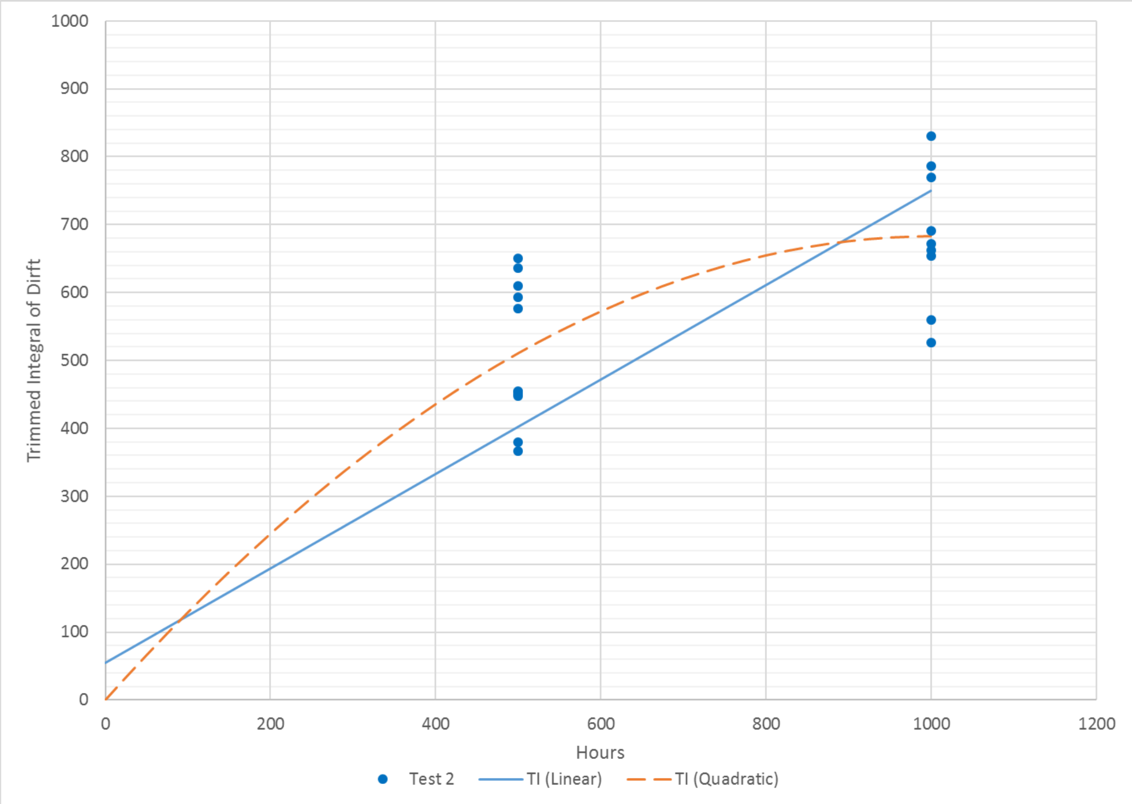


Figure 105: Linear regression model fitting to experimental results (Test 2)

Model	**Equation** (Drift (TI) as a function of Hours (Hrs))	**R-Squared**	**Standard error**
Linear	$TI = 54.32 + 0.6956 \times (Hrs)$	86.0%	116.22
Quadratic	$TI = 0 + 1.359 \times (Hrs) \; - 0.000676 \; \times (Hrs)^2$	93.2%	82.33

Table 23: Error for linear regression models fitted to experimental data (Test 2)

8.2.2 Developing the low hour behaviour of the model

Considering then test 5, with an additional performance inspection completed at 58 hours, the same quadratic regression model can be shown to be a less suitable model for describing the early life regression of TI drift, with underestimation resulting in a higher standard error and R-Squared, as visualised in Figure 106, and presented in Table 24.

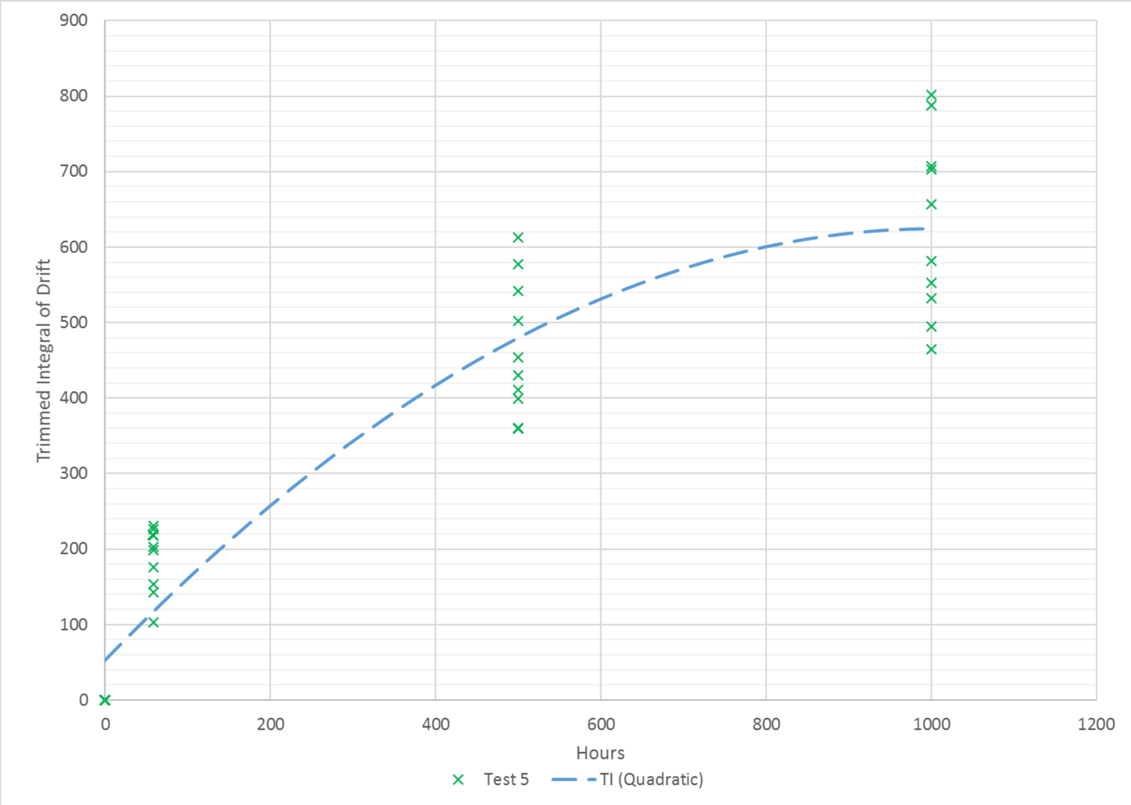


Figure 106: Regression model fitting to experimental results (Test 5)

Model	**Equation** (Drift (TI) as a function of Hours (Hrs))	**R-Squared**	**Standard error**
Quadratic	$TI = 5297 + 1.137 \times (Hrs) \; - 0.000565 \; \times (Hrs)^2$	88.79%	88.90

Table 24: Error for regression model fitted to experimental data (Test 5)

With the additional performance data available at 58hours, test 5 was then used to develop a more suitable non-linear regression model that demonstrated good statistical correlation with the data, while adequately representing limit behaviour. Figure 107 visualises the resultant non-linear regression models, alongside the test data and the pre-existing quadratic model. Table 25 then presents the equation associated with each model, alongside the standard error associated with the correlation of each model.

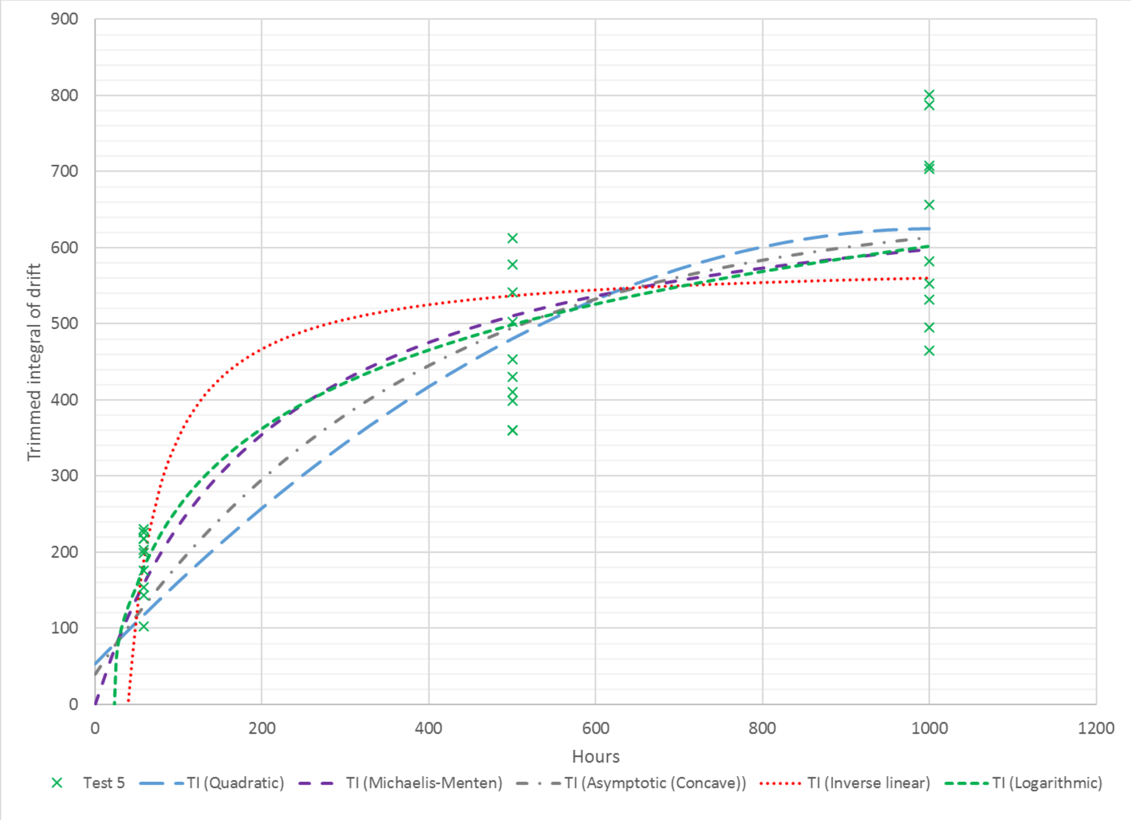


Figure 107: Non-linear regression model fitting to experimental data (Test 5)

Model	**Equation** (Drift (TI) as a function of Hours (Hrs))	**Standard error**
Quadratic	$TI = 52.97 + 1137 \times (Hrs) \ - 0.000565 \ \times (Hrs)^2$	88.9
Michaelis-Menten	$TI = 720.877 * \left(\frac{Hrs}{206.719 + Hrs}\right)$	81.1
Asymptotic (Concave)	$TI = 655.617 - 616.192 \times e^{(-0.00268619 \times Hrs)}$	85.9
Inverse linear	$TI = 583.7 - \frac{23268}{Hrs}$	106.4
Logarithmic	$TI = 148.809 \times \ln(Hrs) - 426.621$	92.06

Table 25: Error for non-linear regression models fitted to experimental data (Test 5)

The lowest standard error is associated with the Michaelis-Menten model, which can also be seen to provide a visually ‘good’ fit to the data. The Asymptotic model also demonstrates a lower standard error than the quadratic model, describing a less steep increase in drift at early hours compared to the Michalis-Menten model. The Inverse Linear and Logarithmic models each result in higher standard errors, likely constrained by the fact that both models pass below the origin. The Inverse linear model represents the steepest slope for increasing drift at low hours, approaching its asymptote earlier relative to the alternative models.

Examining the limit behaviour of each model as it tends to large values of hours run also provides further insight into the potential suitability of each regression model. As can be seen in Figure 108, the quadratic model for TI drift tends to negative values after 2000hours, and the model for logarithmic regression tends to a high value of TI drift as hours approaches infinity, corresponding to the TI drift of ~1000-1100 at the end of vehicle life. The other models all tend to asymptotes, some of which can be observed from the constants in their equations. The Inverse linear model tends to an asymptote at a value of 583.7, which is lower than the mean of TI drift observed at 1000hours of ~628. The Michaelis-Menten and concave asymptotic models tend to asymptotes between ~650 and 720, which compare more favourably to the mean observed at 1000hours. All available empirical data and expert judgment suggest that an asymptotic regression is most suitable for describing both NCV seat wear and injector performance drift.

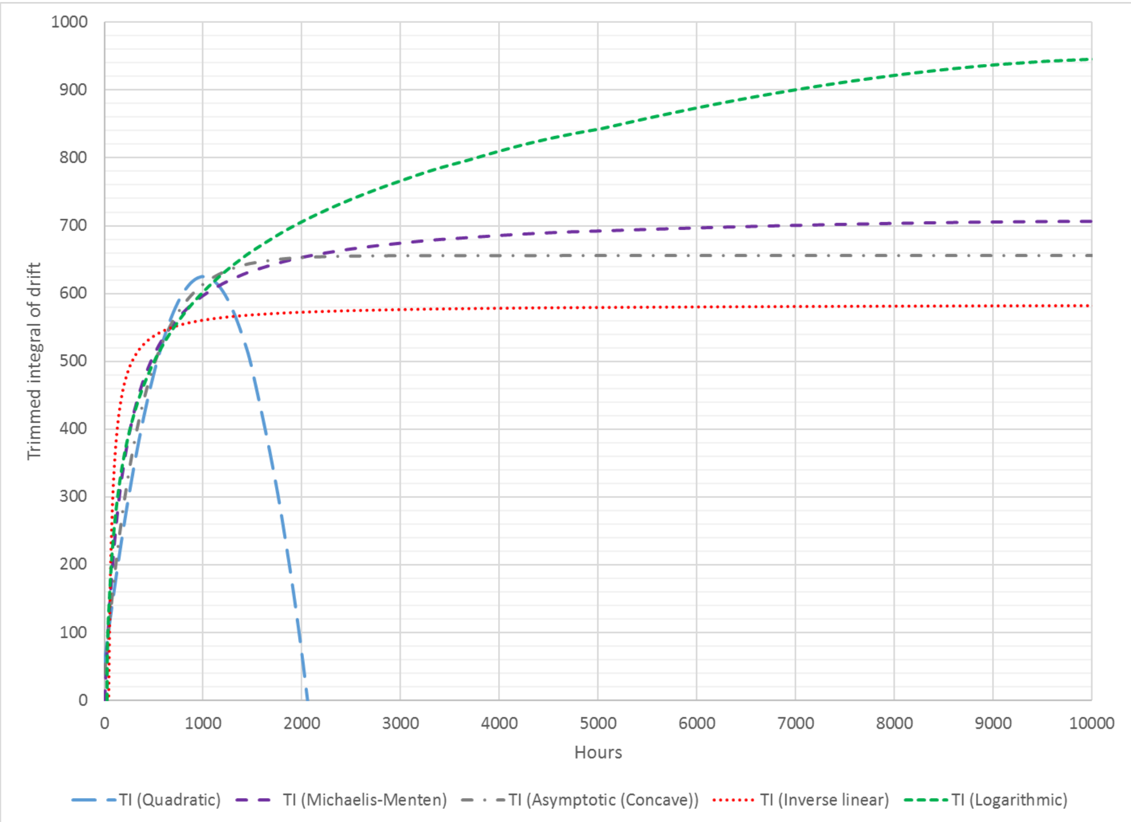


Figure 108: Limit behaviour of non-linear regression models

Of the regression models presented, the Michaelis-Menten model represents the best statistical correlation to the data, while matching the behaviour of the working mental model of the failure mode. However, from the data available, the Michalis-Menten model can be shown to under-estimate the magnitude of TI Drift at higher hours. Figure 109 shows the regression model fitted to the data from test 5 plotted against the sample mean for each performance inspection interval. It is anticipated that with additional data points from empirical testing, the model could be further refined to best fit the relationship between TI Drift and hours run, but the model represents a working approximation of the failure mode.

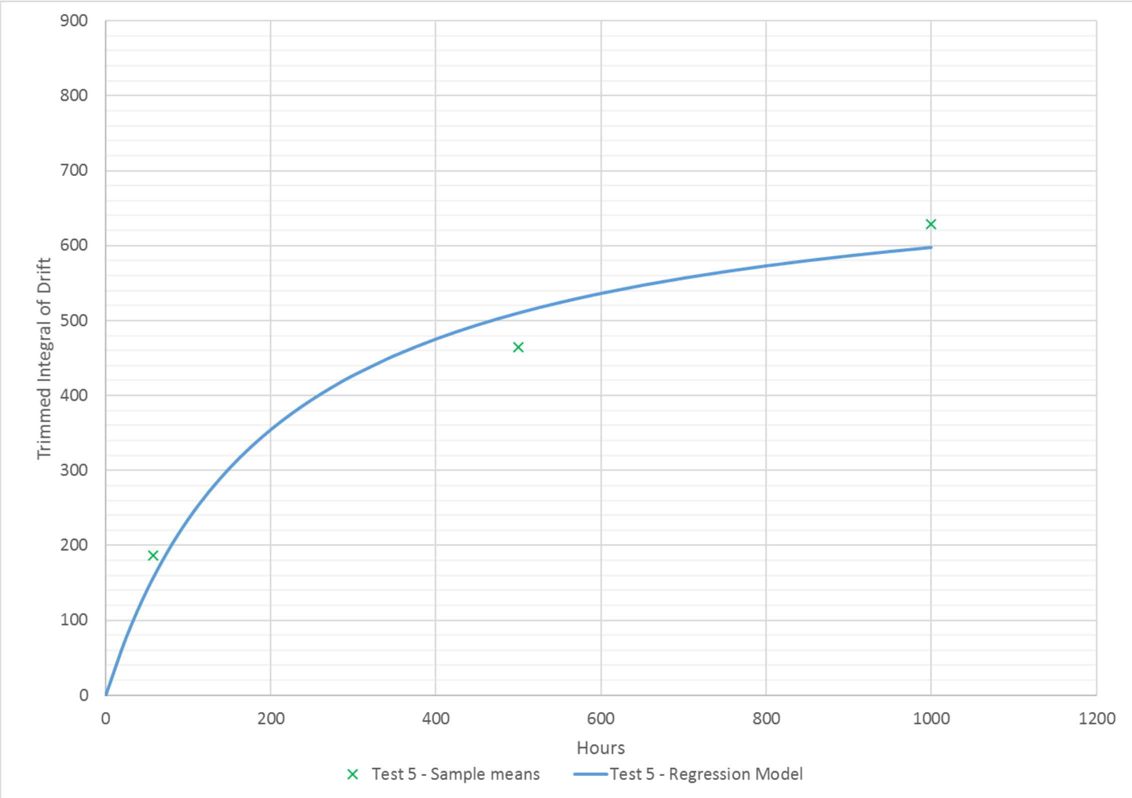


Figure 109: Fitted Michaelis-Menten model plotted against sample means for Test 5

8.2.3 Developing a generalised model of drift over time

Having establish that the Michaelis-Menten regression model most adequately describes drift over time behaviour when low hour results are included, the next step in developing a generalised model of drift over time is to fit it to the results of the other tests. Re-considering test 2, Figure 110 visualises the results of test 2, with both the previous quadratic model and the Michaelis-Menten regression models fitted, with the results and fitted model for test 5 shown for reference.

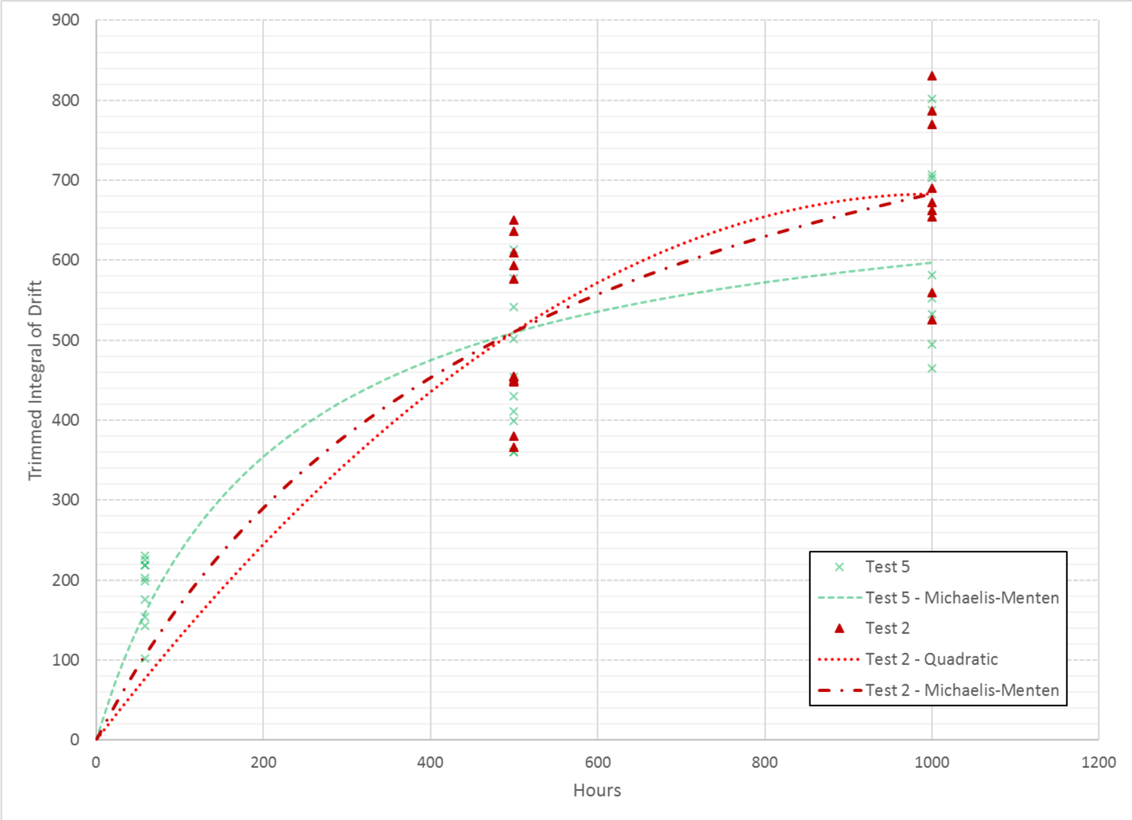


Figure 110: Application of Michaelis-Menten model to additional experimental results (Test 2)

While the Michaelis-Menten model can be seen to be a good correlation to the available data, its behaviour at low hour may not be adequate given the results associate with test 5. Given that the mean TI Drift for test 2 at both 500 and 1000 hours were greater than those associated with test 5, it follows that the regression model associated with test 2 would result in a greater value of TI drift for every positive value of hours. However, the Michaelis-Menten model as fitted demonstrates lower values of TI drift until 500 hours.

In its general form, the Michaelis-Menten model is shown in Figure 111 and can be described by the Equation 7.

$$f(x) = Theta1 \times \frac{x}{Theta2 + x} \tag{7}$$

Theta 1 defines the asymptote, and Theta 2 corresponds to the value of x such as to satisfy Equation 8.

$$f(Theta2) = Theta1 \times \frac{1}{2} \tag{8}$$

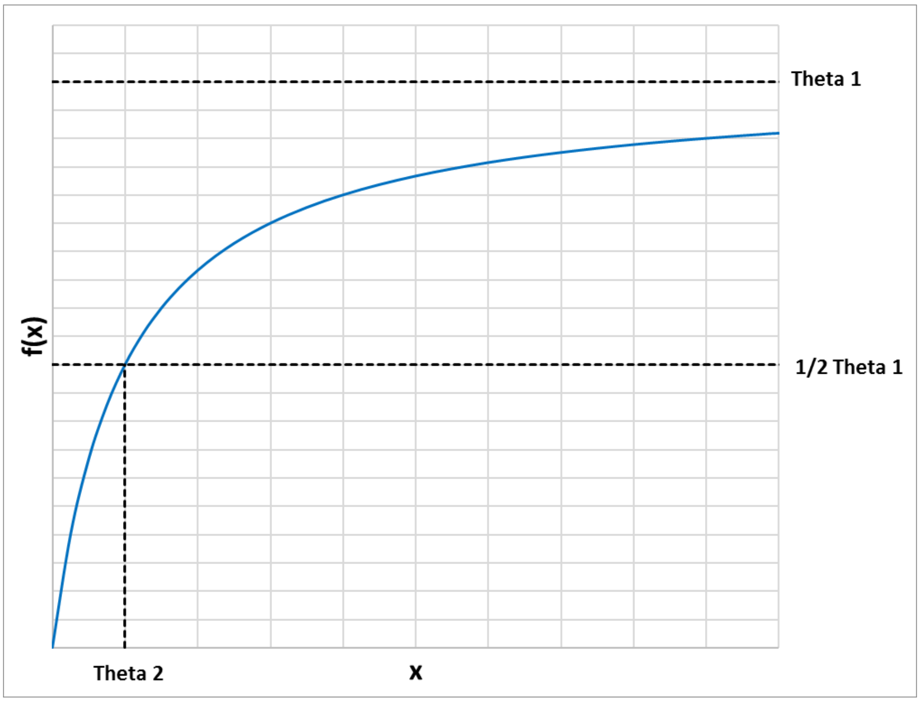


Figure 111: Generalised form of Michaelis-Menten regression model

It can therefore be inferred that the population means of TI drift at high hours are going to have more influence on Theta 1, while Theta 2 will be influenced by any available data at lower hours. In the case of test 2, and more generally for the original DoE tests, there is no such data available at lower hours, hence why the fitted model can be perceived to fit inadequately between 0 and 500 hours. It is therefore proposed to place a constraint on the value of Theta 2 when performing regression analysis, and that value should be related to that observed for test 5 where appropriate data was available.

Applying an upper constraint on Theta 2 of 200, compared to a value of 206 observed for test 5, and re-performing regression analysis, a constrained model to describe TI drift over time for test 2 can be developed, as shown in Figure 112, while the equations and standard errors are shown in Table 26. While the resultant model results in a higher standard error for the available empirical data, it represents a better approximation of the shape of the regression curve given the data

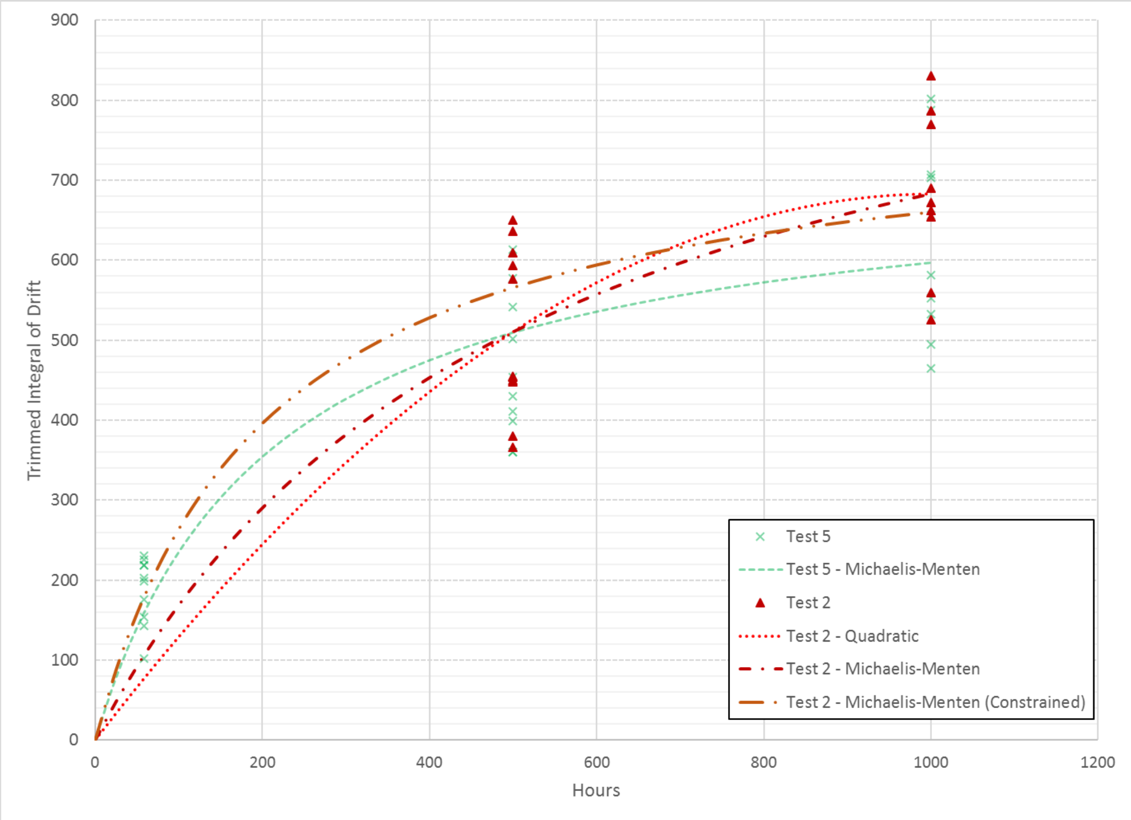


Figure 112: Visualisation of the effects of applying a constraint on Theta 2 for the model as fitted to Test 2

Model	**Equation** (Drift (TI) as a function of Hours (Hrs))	**Standard error**
Quadratic	$TI = 0 + 1.359 \times (Hrs) \; - 0.000676 \; \times (Hrs)^2$	82.33
Michaelis-Menten	$TI = 1032.89 * \left(\frac{Hrs}{511.154 + Hrs}\right)$	80.89
Michaelis-Menten (Constrained)	$TI = 791.665 * \left(\frac{Hrs}{200 + Hrs}\right)$	88.63

Table 26: Error for Michaelis-Menten model with and without constraint on Theta 2

Applying the same regression model, with constraints as appropriate, a series of regression curves can then be generated for each test associated with the DoE. The results are visualised in Figure 113, with the equations presented in Table 27.

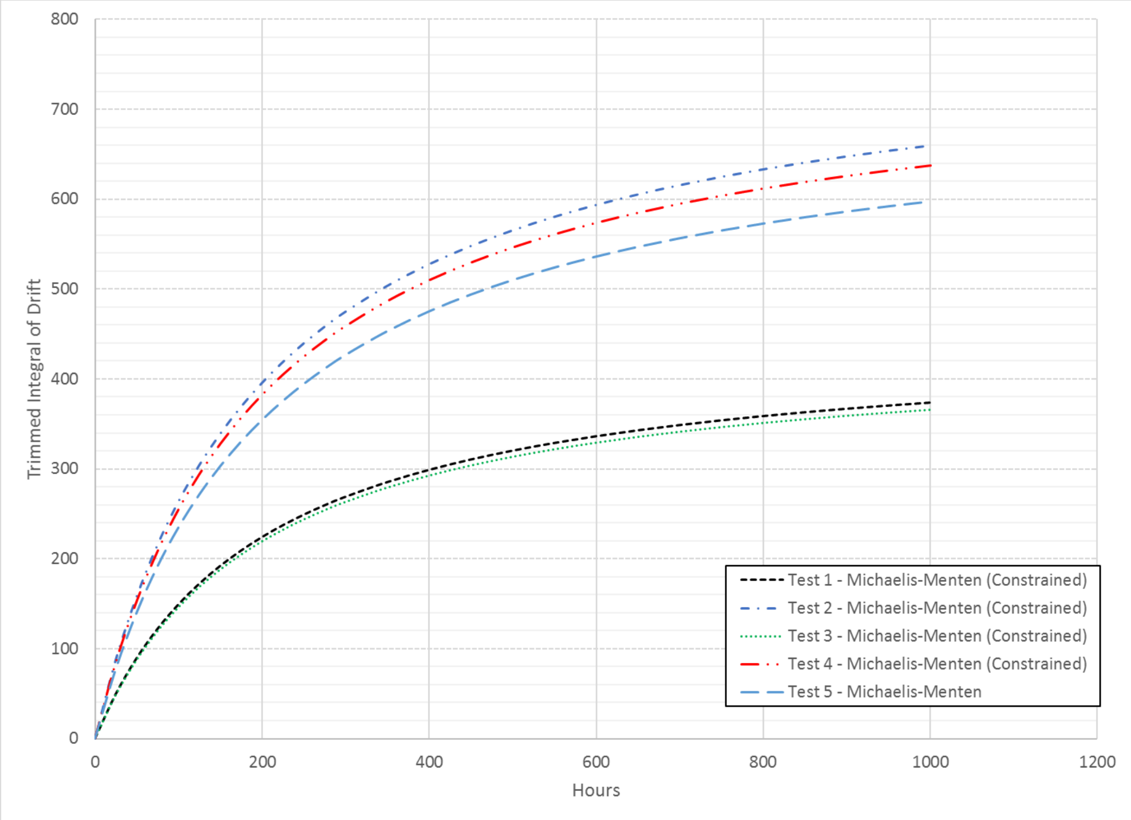


Figure 113: Michaelis-Menten models, with constraint on Theta 2, as fitted to Tests 1-5

Test	**Equation** (Drift (TI) as a function of Hours (Hrs))	**Rail Pressure**	**Sample Mean**
1	$TI = 448.837 * \left(\frac{Hrs}{200 + Hrs}\right)$	1800	332
2	$TI = 791.665 * \left(\frac{Hrs}{200 + Hrs}\right)$	2500	722
3	$TI = 438.608 * \left(\frac{Hrs}{200 + Hrs}\right)$	1800	314
4	$TI = 764.635 * \left(\frac{Hrs}{200 + Hrs}\right)$	2500	649
5	$TI = 720.877 * \left(\frac{Hrs}{206.719 + Hrs}\right)$	2150	628

Table 27: Regression equation, rail pressure, and sample mean for each test

From observing the approximated regression equations and considering both the Rail Pressures and sample means associated with each test at 1000hours, further inferences can be made with reference to a generalised regression model for TI drift. Figure 114 visualises the approximated value of Theta 1 for the regression model associated with each test, plotted against the associated sample mean. A linear regression line has then been fitted to approximate the relationship, with an R^2 of 99.45%, and a standard error of 54.2, described by Equation 9.

$$Theta\ 1 = 1.1686 * (Sample\ mean\ of\ TI\ Drift\ @\ 1000hours) \tag{9}$$

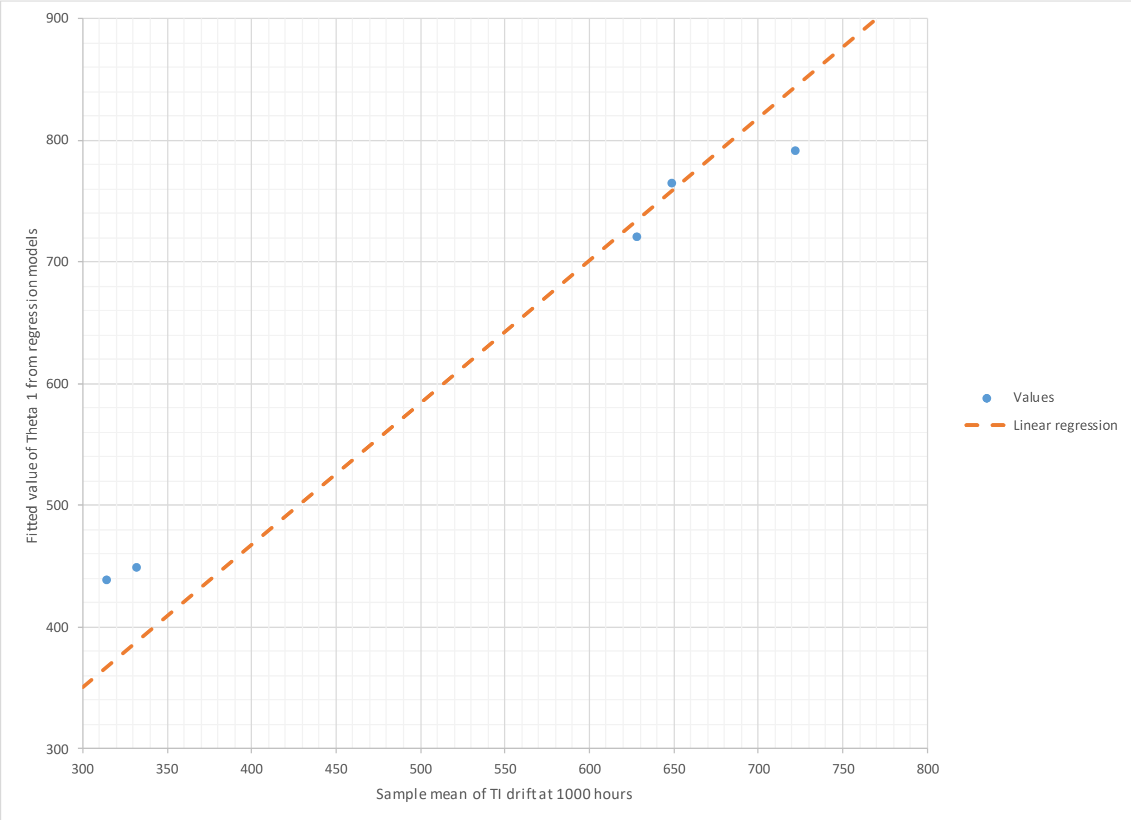


Figure 114: Regression model for sample mean and fitted Theta 1

Note that an alternative linear regression model, including an offset to the origin can be determined that results in a better statistical correlation, with a standard error of 16.5. However, with limited data available for regression, and with a view to avoiding overfitting the data when no such offset is anticipated, a simple proportionality is assumed at this time.

It has therefore been shown that for a given mean value of TI drift at 1000hours, the value of Theta 1, describing the asymptotic value, can be approximated using a linear relationship. In §7.11, a generalised model of mean TI drift for a given pressure at 1000 hours was presented in Equation 5. Combining that model, with the approximation of Theta 1 gives Equation 10.

$$Theta\ 1 = 1.1686 * \left(314 + 454.799 \times \frac{(Pressure - 1800)}{157.047 + (Pressure - 1800)}\right) \tag{10}$$

If Theta 2 is assumed to be approximately equal to 200 for all pressures, a generalised model of TI drift for a given number of hours, at a given rail pressure, can then be derived as shown in Equation 11.

$$TI = \left(1.1686 * \left(314 + 454.799 \times \frac{(Pressure - 1800)}{157.047 + (Pressure - 1800)}\right)\right) \times \frac{Hours}{200 + Hours} \tag{11}$$

Figure 115 then presents a visualisation of this generalised model for multiple values of Rail Pressure within the boundaries of the experimental space.

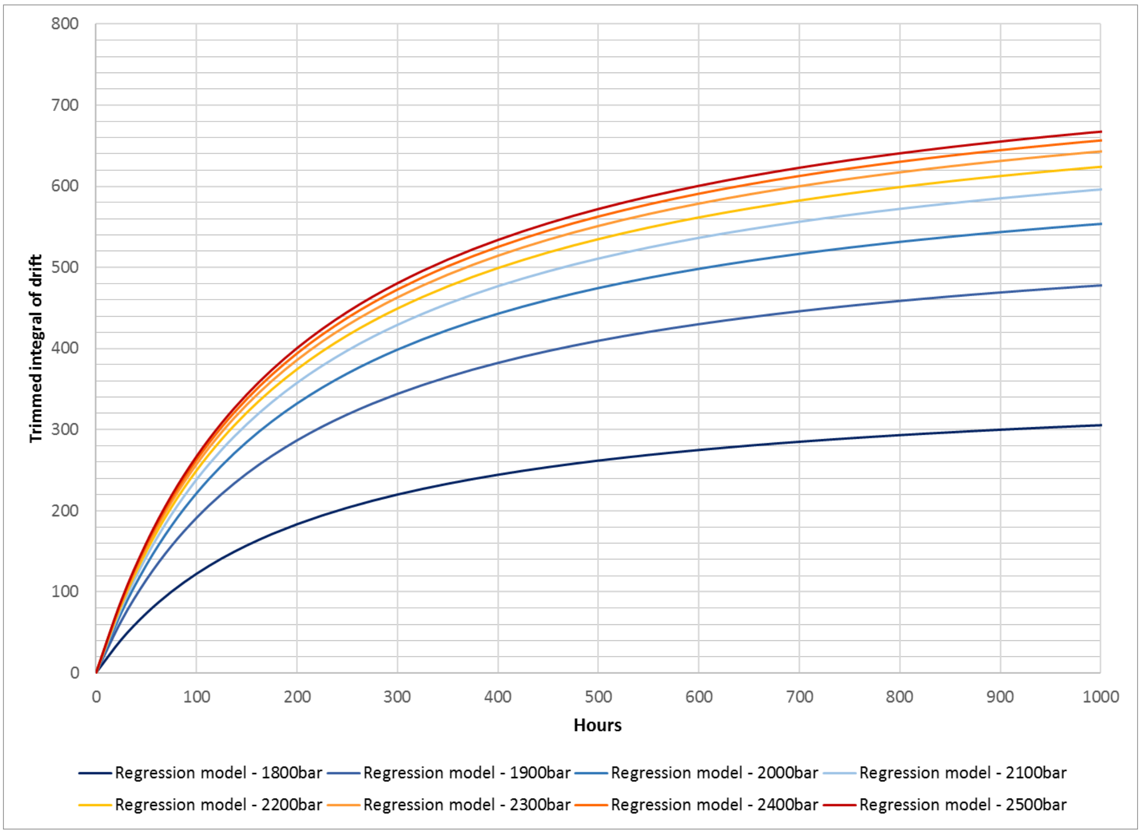


Figure 115: Generalised regression model for TI drift progression

8.3 Validating the generalised model of drift with pre-existing data

One of the original motivations for this empirical study was to determine if there was a difference in the drift over time behaviour of samples tested to hydraulic rigs compared to engines and vehicles, and if so, to explain that difference using an empirically based model.

To begin to do so, the results for samples testing on hydraulic rigs were compared to the results of this research. A sample of results were selected from a larger population at random, with design and build specifications equivalent to the samples used for this test, and all tested on equivalent test cycles. Samples from a total of 9 tests were included, representing a sample population of 50 injectors[3]. Except for one test, each sample group had multiple performance inspection intervals associated with it between 100 hours of test time, and 2500 hours. The performance inspection results were re-processed for TI drift in the same manner as used for this research. Figure 116 visualises the results as grouped by test.

[3] This FIS product is used on applications with both 4- and 6-cylinder engines, hence why the sample population is not a multiple of 6

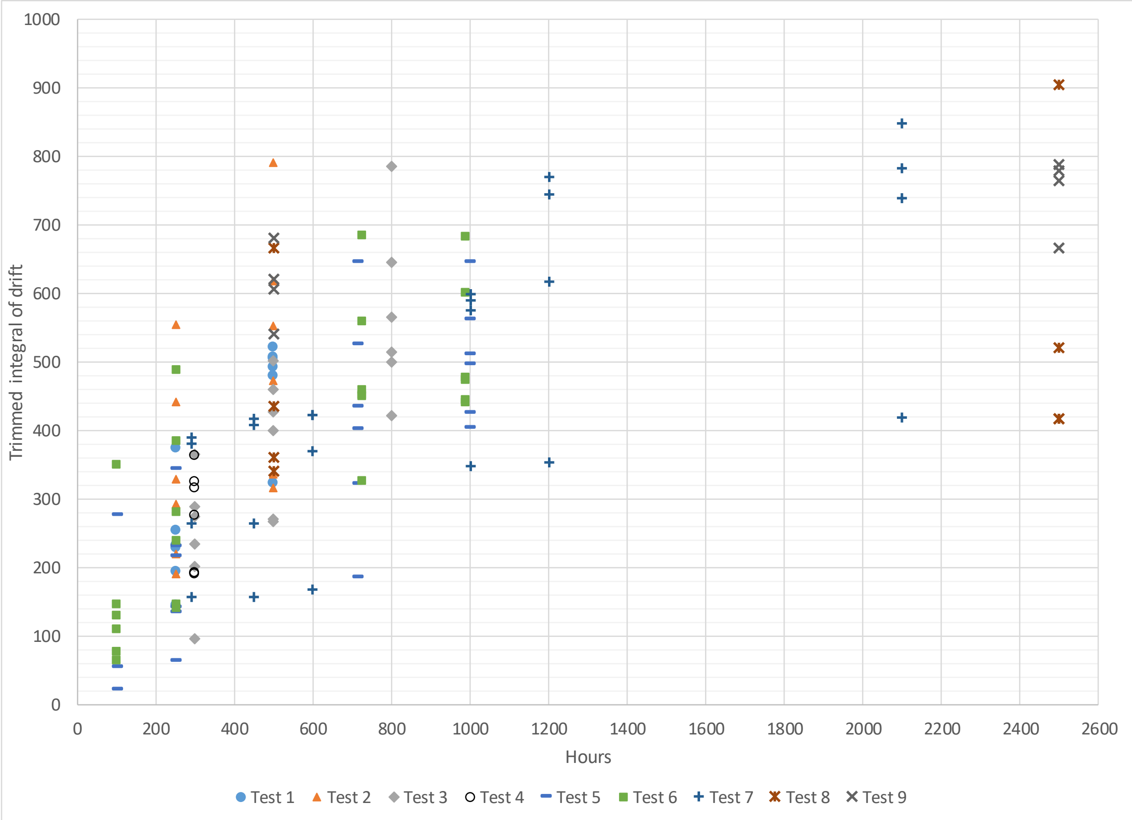


Figure 116: Visualisation of TI drift progression over time for prior samples

In observing the data as either multiple smaller populations, or a singular larger population, a similar asymptotic shape can be inferred as seen for the results associated directly with this research. Simplifying this data as a single population of injectors of the same design and build, tested on the same test cycle with the same fuel, and a regression can be developed in a similar manner to that presented previously based on a large sample size. Using the Michaelis-Menten model, and similar constraints on shape as used previously, a regression model can be developed as shown in Figure 117.

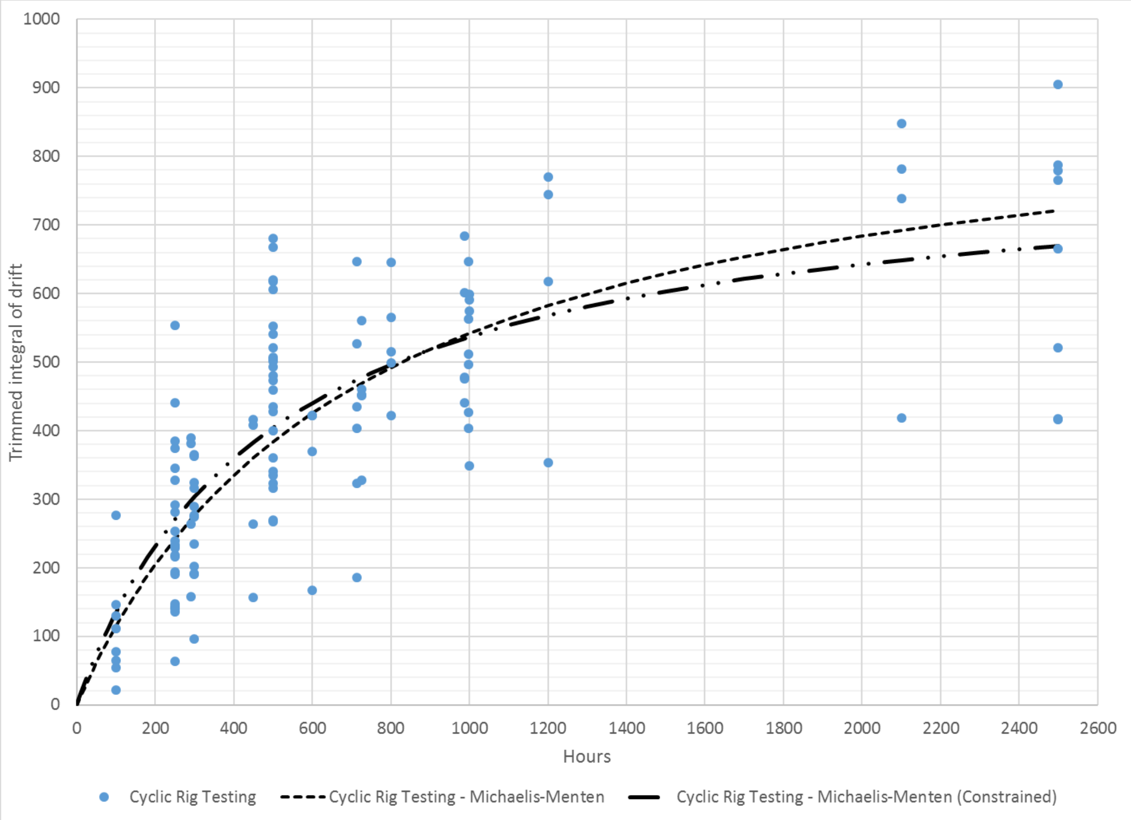


Figure 117: Michaelis-Menten model fitted to prior data with and without constraint on Theta 2

Figure 118 then compares that constrained model for TI drift for cyclic rig testing with the generalised model of TI drift presented in §8.2.3. As can be seen, despite the cycle including a maximum rail pressure of 2400bar, the model for cyclic rig testing only reaches a similar level of drift after 1000hours as expected for a rail pressure of ~2000bar, and the shape of the curve at low hours is significantly less steep. As such, it can be said that the steady state rig testing represents an acceleration with respect to cyclic testing for TI drift.

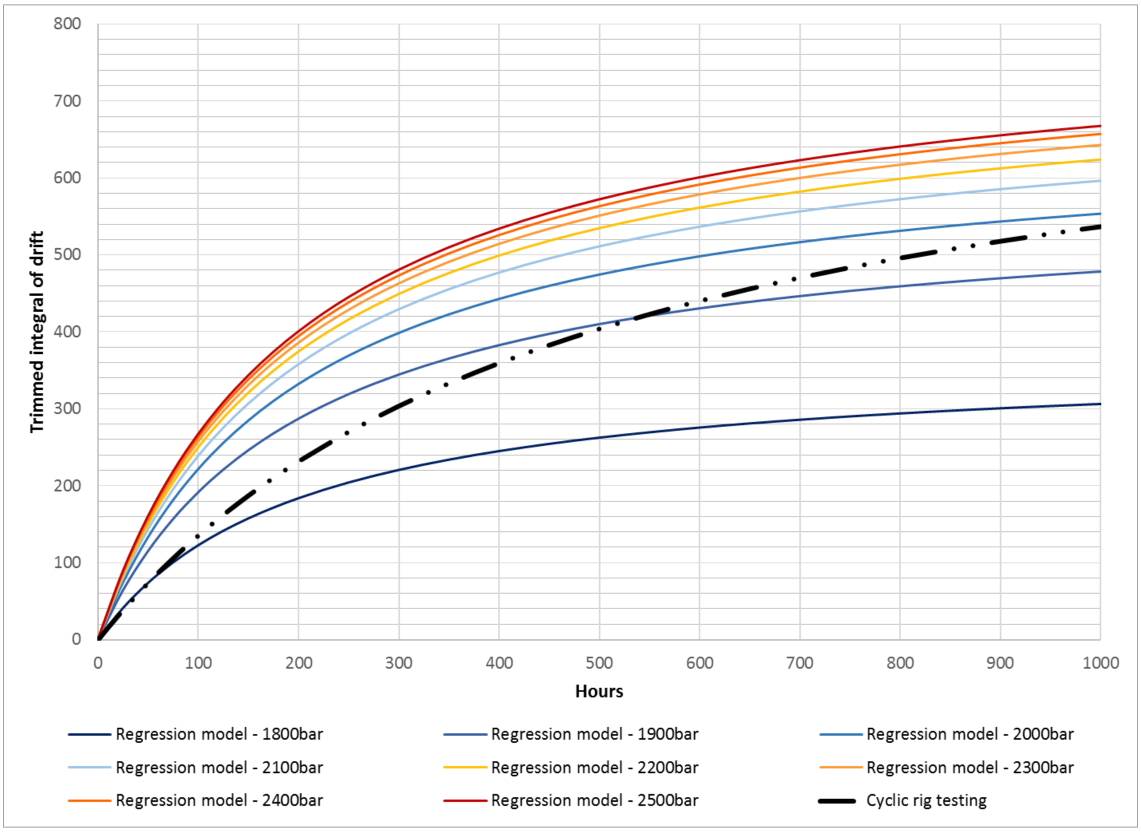


Figure 118: Comparison between regression model for prior data and generalised model of TI drift progression

In order to further understand both the relationship between TI drift and different pressures, and the relative accelerations associated with cyclic and steady state testing, the test cycle, in terms of pressure residencies, needs to be explored further. Table 28 presents the test cycle, in terms of steady state points described by rail pressure and time durations, along with summary statistics. As can be seen, the average pressure of the 270 second cycle equates to <1600bar, while one third of the cycle duration is spent at a pressure of 2400bar.

Time (s)	Pressure (Bar)
30	1400
45	2400
45	2400
30	1700
30	900
30	1600
30	1000
30	500

Total Duration	270
Average Rail Pressure (w.r.t. time)	1589
Residency @ 2400bar	33.3%

Table 28: Test cycle associated with prior data

When considering the regression model for cyclic testing and the generalised pressure-based model, the TI drift associated with cyclic testing is greater than that that would be expected for the average pressure of the cycle. However, if the test cycle is considered based on percentage duration spent at rail pressures within the boundary of the experimental model, which for this cycle is only the cycle points at 2400bar, then a time-based transformation may be applied to the data. If a 33% time transformation is applied to the results of the cyclic rig testing and a regression model is re-calculated using the Michaelis-Menten model, the results can again be compared to the generalised model. Figure 119 visualises the regression model for cyclic rig testing after a time-based transformation had been applied to the results. As can be seen, with the time-based transformation, and considering a pressure of 2400bar, the regression model for cyclic rig testing is much closer to the equivalent pressure level for the generalised model of drift.

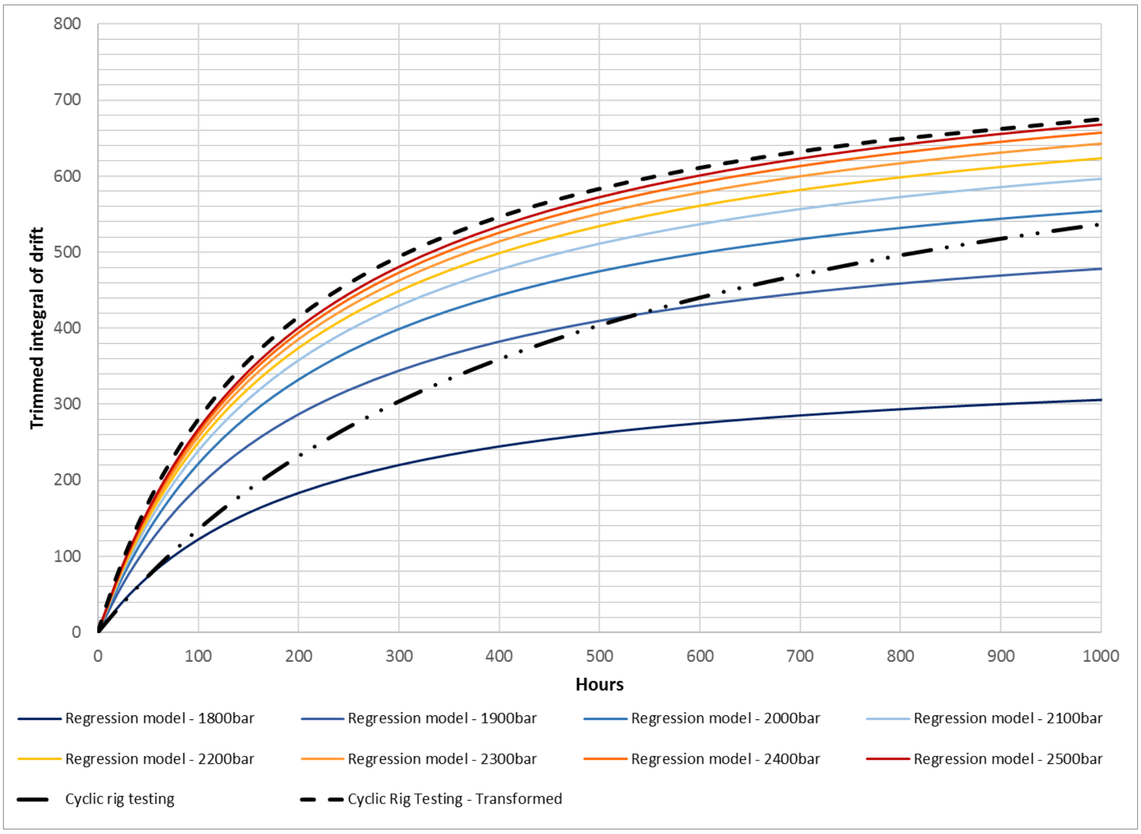


Figure 119: Visualisation of effects of the application of a time-based transformation to prior data

From this analysis, it can be suggested that average pressure of a test cycle, or drive cycle, is not a suitable means for then predicting the levels of injector drift, and it is instead the proportion of time spent at higher pressures that allows a more accurate approximation to be made.

8.4 Using the generalised model of drift to interpret the potential differences between applications

8.4.1 Using the generalised model of drift to compare different vehicle duty cycles

Using the generalised model of drift presented in §8.2.3, and the knowledge that the proportion of time spent at high pressures that provides the most suitable approximation of injector drift, different vehicle duty cycles can be compared against each other, and against testing on hydraulic rigs to determine relative accelerations.

Suitable application data is available for vehicle development as supplied by DPEs customer for this FIS system, using a combination of ETC monitoring, GPS systems, and additional data measurement. In this case, the data was captured from the same vehicle and powertrain combination, so the data represents differences in drive environment and use case only. The 3 different drive types used for this comparison are as visualised in Figure 120, and each data capture represents approximately one hour of driving. The 'Highway' drive type represents a point to point journey along a multilane road with high drive speeds, but low engine load requirements, typical of a haulage application. The 'Urban/Extra-Urban' drive type then represents a journey through a town or city, with a mixture of single and multiple lane roads, multiple acceleration and deceleration events, and a mixture of drive speeds. The 'severe' drive type then represents a drive up a mountain road, where drive speeds are low, and high engine loads are required.

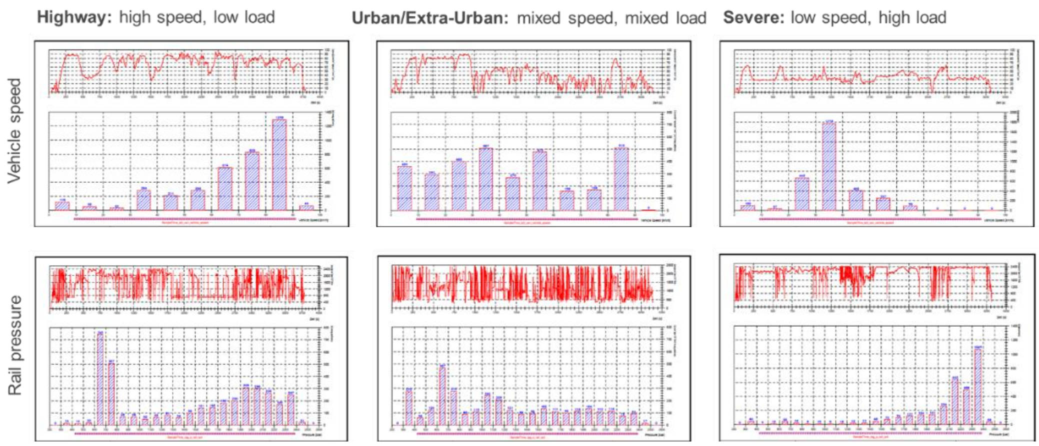


Figure 120: Visualisation of 3 different vehicle duty cycles in speed and load collectives

From the data captured, it is then possible to compare the different drive types with regards to distributions and averages. Figure 121 compares the cumulative pressure residency, in terms of drive time, of the 3 different drive types. As can be seen, the highway drive type has a high proportion of time spent at rail pressures between 600-800bar, corresponding to the low engine loads associated with cruising. The Urban/Extra-Urban drive type has a more linear distribution of rail pressures, including almost 10% of the time at pressures of ~400bar, which corresponds to the engine idling when the vehicle is stopped owing to road conditions or traffic management. The severe drive type then can be seen to spend the majority of the drive time at the highest pressures.

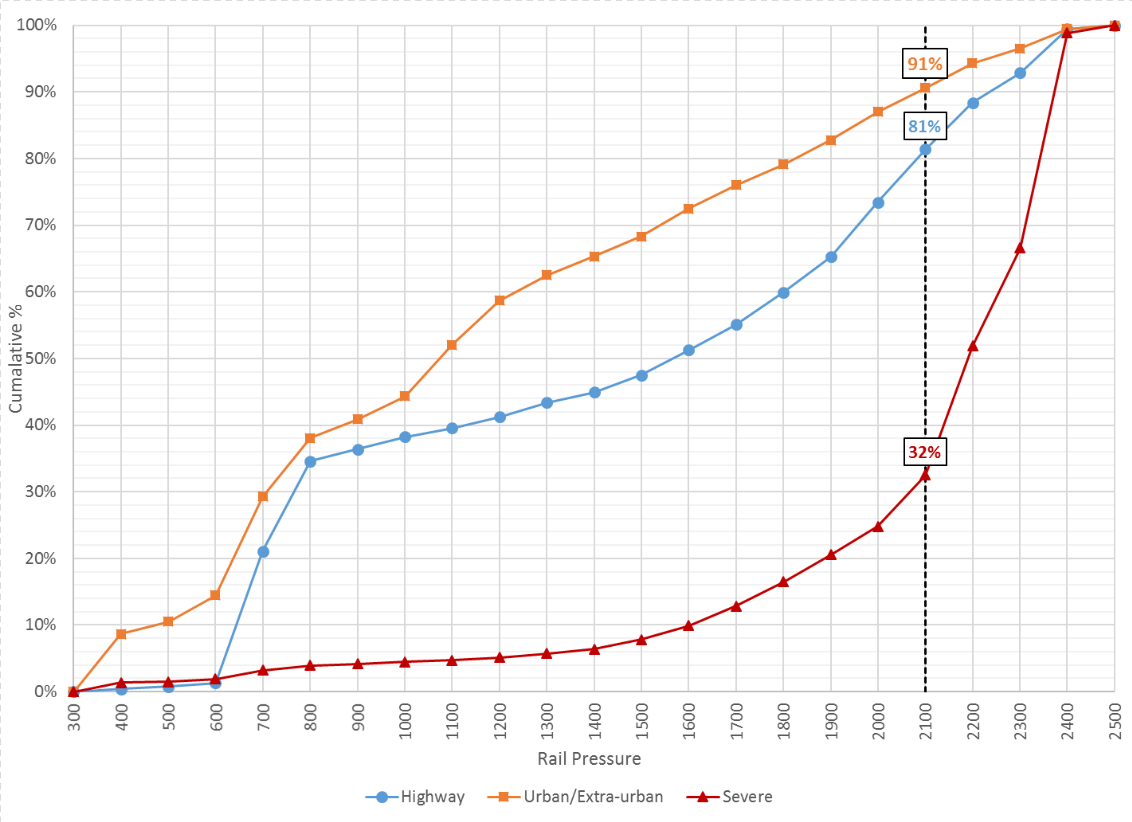


Figure 121: Vehicle duty cycles expressed as cumulative pressure residencies

The average vehicle speed for each drive type, alongside the proportion of drive time where the rail pressure is greater than 2100bar, is then summarised in Table 29. As can be seen, as expected, the highway drive type has a significantly higher average vehicle speed, while the severe drive type has the lowest average speed, despite not featuring frequent stoppages. Pressure residency over 2100bar was selected as it is a rail pressure central to the range modelled in the generalised model of drift, and as a result of the probable non-linearity in drift with rail pressure, represents potentially significant levels of drift. As can be seen, the Urban/Extra-urban drive type only corresponds to less than 10% of the drive time at pressures greater than 2100bar, while the severe drive type corresponds to over 65%.

	Highway	Urban/Extra-urban	Severe
Average Speed	67.5	44.2	35.4
Time residency @ >2100bar	18.6%	9.4%	67.5%

Table 29: Comparison of applications in terms of speed and residency at high rail pressure

Using the generalised model for TI drift, and the residency of each drive type over 2100bar, the TI drift for hours run for each drive type can be approximated. Figure 122 visualises the relationship between driving time and TI drift for each drive type, and it can be seen that the Severe drive type represents a >5x acceleration compared to the Urban/Extra-urban drive type at a nominal drift level of 400, while the Highway drive type represents only a ~2x acceleration.

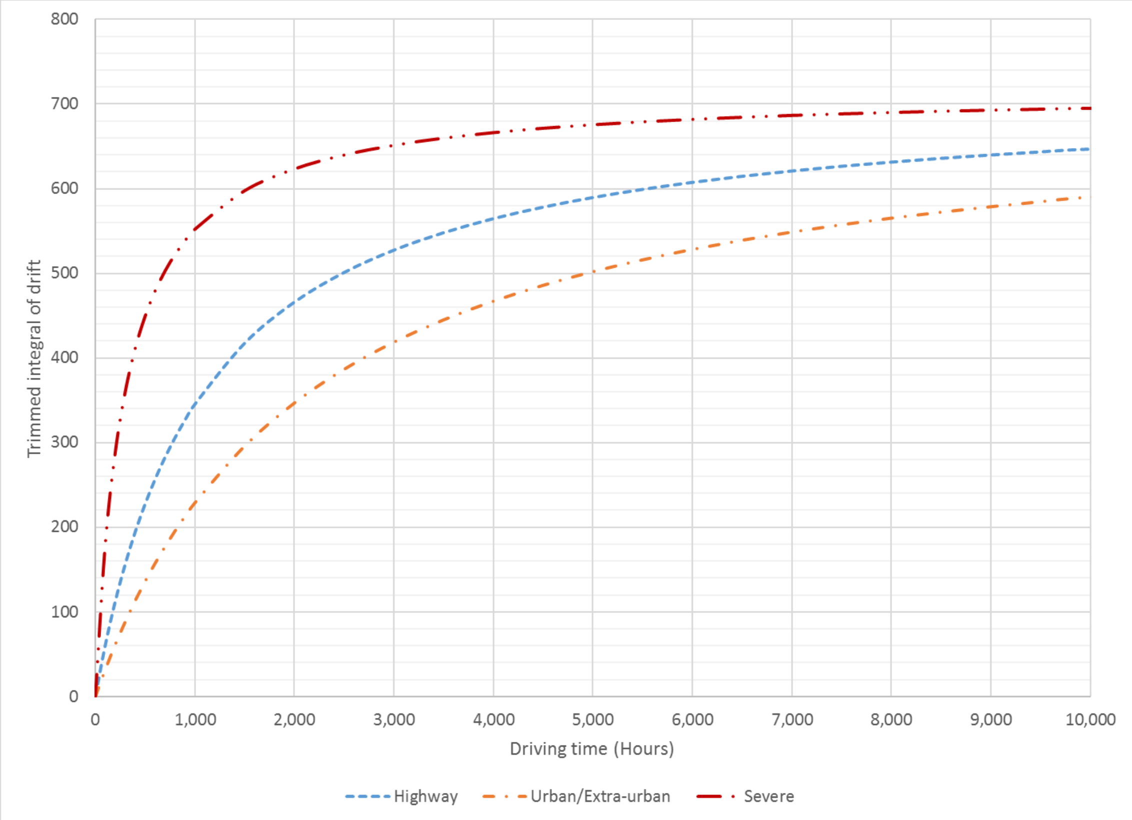


Figure 122: Comparison of different drive types using the generalised model

Since warranty is based on vehicle mileage rather than engine hours, the different drive types can instead be considered in terms of the relationship between TI drift and driving distance. By multiplying hours by the average vehicle speeds associated with each drive type, TI drift can be expressed in terms of approximate driving distance, as visualised in Figure 123. As can be seen, with respect to vehicle mileage, the low vehicle speeds, coupled with the high pressure residencies of the severe drive type represents an increased acceleration of ~10x compared to the urban/extra-urban drive type at a nominal drift level of 400, while the increased vehicle speeds associated with the highway drive type reduce its acceleration with respect to the urban/extra-urban drive type for engine hours.

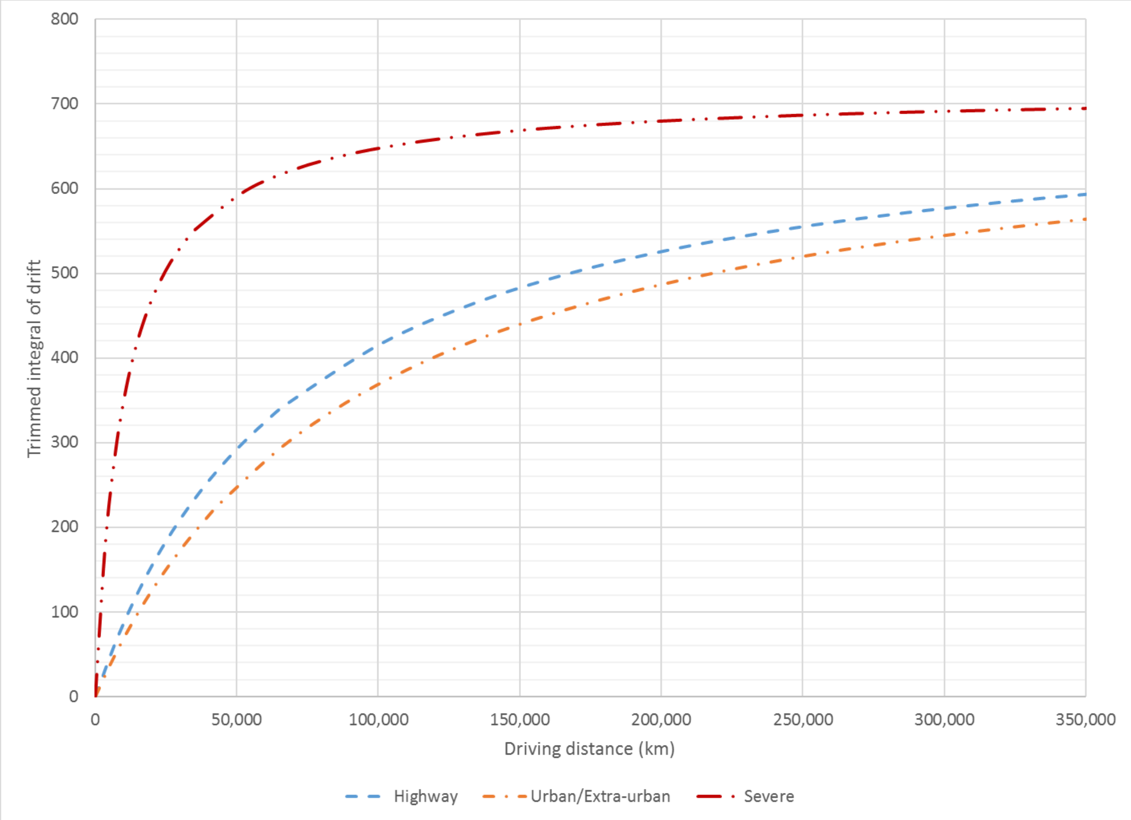


Figure 123: Comparison of different drive types with a time-based transformation applied

8.4.2 Using the generalised model to compare TI drift associated with different test cycles and vehicle usage

The progression of TI drift for different test cycles and application drive types can be compared to assess the relative accelerations associated with each. Figure 124 visualises the TI drift for both cyclic, and steady state hydraulic rig testing, alongside the 3 drive types discussed in §8.4.1, in terms of hours run. In the case of the two hydraulic rig tests, the cyclic rig data is fitted to empirical results with no time-based transformation applied, while the steady state test represents testing at a rail pressure of 2500bar, as predicted by the generalised model of TI drift presented in section §8.2.3. As can be seen, the cyclic rig testing does not represent a significant acceleration over the most extreme of the drive types until over 1400hours are reach, while testing at steady state represents a ~2x acceleration at a nominal TI drift level of 400.

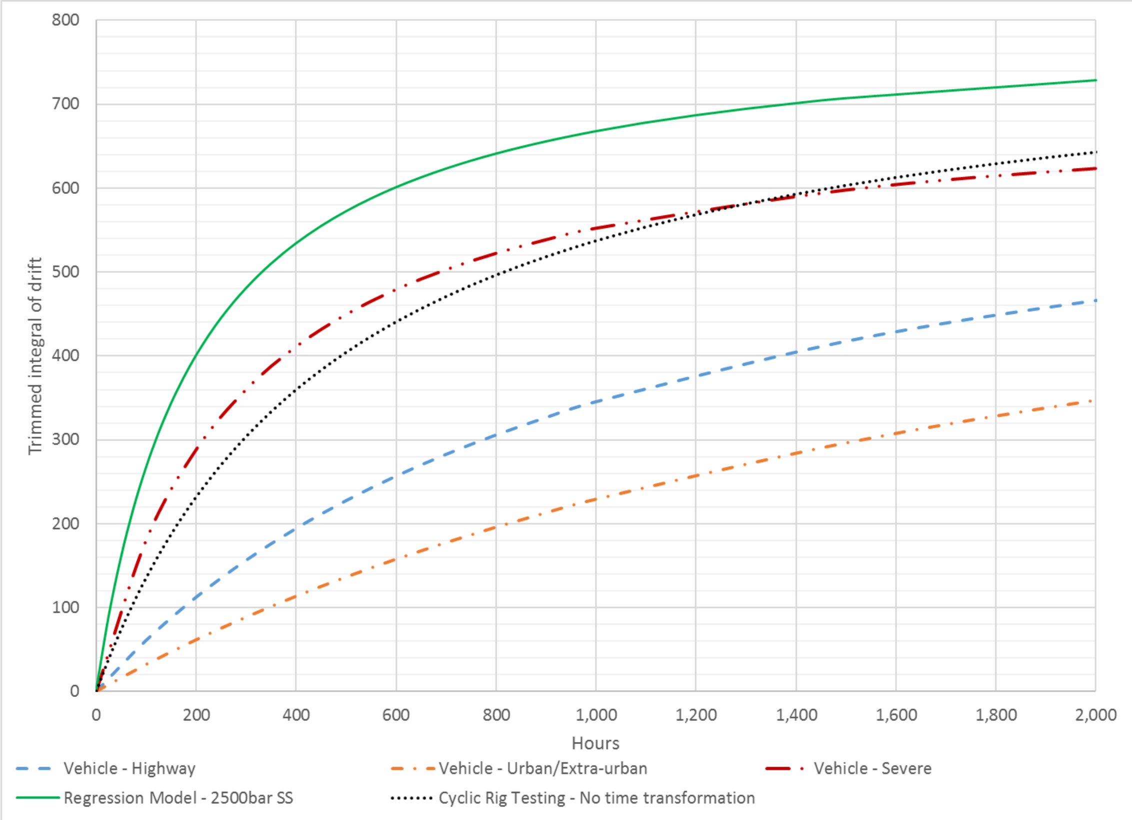


Figure 124: Comparison of models of drift for hydraulic rig testing and drive types

However, when EU driving laws are considered, which limit total drive time to 56 hours per week, or 33%, the acceleration associated with hydraulic test rigs should be refined further. Delphi Technologies validation plans typically assume that a hydraulic test rig will be running for ~70% of the 168 hours in a week. This considers stoppages associated with the test plan and FIS, refuelling and other fluid servicing of the test rig, and stoppages to work on any adjacent test rigs. With an upper limit of 33.3%, and 70% applied accordingly, the models can be re-calculated, as visualised in Figure 125, where total elapsed hours are plotted on the x-axis rather than hours of running, and the utilisation percentage for each case is annotated on the right hand side. As can be seen, when the models are considered in terms of total elapsed hours, both hydraulic rig tests represent meaningful accelerations over even the most severe vehicle drive type.

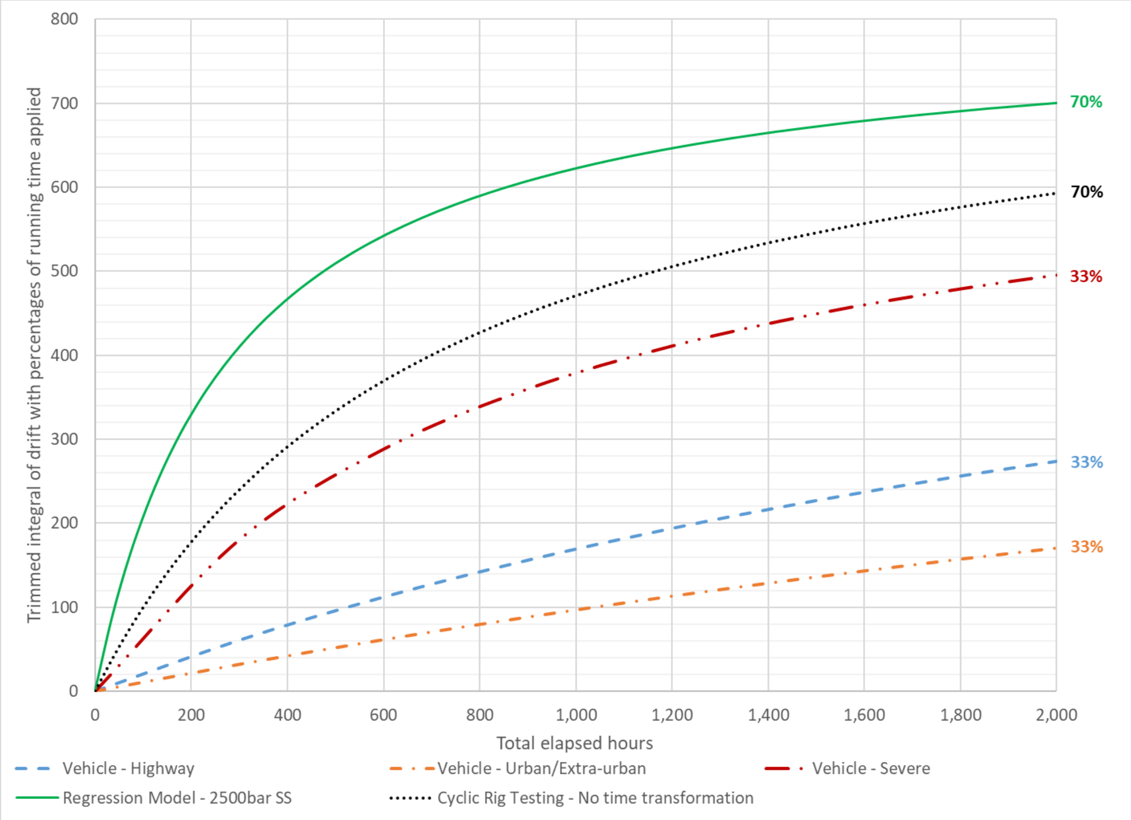


Figure 125: Comparison of models of drift with utilisation percentages applied

8.5 Summary & Conclusions

This Chapter has discussed the final element of the FMC case study associated with this thesis, in which the results of the empirical investigation presented in Chapter 7 are used as the inputs to a regression modelling process. The results of injector performance characterisation tests, in the form of the TI drift metric, along with additional low hour results from Test 5, have been used to describe a series of non-linear regression models using the Michaelis-Menten model. The resulting regression models are asymptotic in nature, and represent a good statistical fit, while corresponding to the working mental models.

A generalised model of injector performance drift is then developed that describes the progression of the failure mode with respect to rail pressure, and duration of operation. That generalised model was then tested against prior empirical results of the product and was demonstrated to represent an adequate fit when a time transformation was applied to the data proportionate to the test cycle residency at higher pressures. Finally, data taken from the application is used to compare different duty cycles, and hydraulic rig testing, using the generalised model, allowing for additional comparisons to be made of relative acceleration factors. In doing so, it can be demonstrated that the duty cycle associated with a specific test cycle or application can have a significant effect on the progression of the failure mode.

This Chapter has demonstrated how deeper understanding of the failure mode is possible through structured further analysis of the empirical results, and through regression modelling, and how that understanding can be validated against prior data. The next Chapter will discuss the design, execution, and results of the case study associated with this research, and how it can be generalised.

Chapter 9 Discussion of Results and generalisation of the method

9.1 Introduction

This Chapter discusses the case study associated with this research, in terms of its design, execution, and results. Reflections on the individual elements of Chapter 6, Chapter 7, and Chapter 8 will be presented, before the gaps in knowledge that remain are discussed. The case study as a whole will then be discussed, reflecting on its suitability as an FMC method, its sustainability, and its potential for generalisation beyond this case study. Finally, metrics suitable for determining the efficacy of an FMC method, as part of an NPD process, are presented.

9.2 Reflections on the Expert Elicitation and System Modelling element of a systematic approach to failure mode characterisation

This section will reflect on the expert elicitation and system modelling element of the FMC method discussed in Chapter 6, both as individual elements, and in terms of their combination.

9.2.1 Reflections on the design, execution, and results of the Delphi Study

The high participation and speed of response submission exhibited in round 1 show both the appetite and general willingness of the expert panel to contribute. The timing of the first round was such that the panel would have the email waiting in their inbox when they arrived at work as to avoid it getting passed over later in the day and by the lunchtime of that day, 3 experts had submitted their responses, and a total of 5 responses had been received by the close of business, representing a 42% response rate after 1 day. This suggested an appropriate panel & subject selection with good initial engagement.

This case study also demonstrated the increased value of group expert elicitation over the expert judgement of a single technical expert. The iterative ranking of the significance of each variable identified by the panel resulted in consensus being reached that a total of 11 variables were of little to no significance to the severity of NCV seat wear. This demonstrates that the combined expert judgement of a group can prove to be better than the judgement of any one if its members.

This case study also highlighted the discord present between the experts distributed around Delphi Technologies on the significance of several variables. 13 variables were identified were informed disagreement was met on their significance, alongside 5 other variables where a small proportion of experts held opinions contrary to the majority of the panel. The former presents areas for future experimental and analytical work, while the latter highlights potential gaps in knowledge sharing and combination within Delphi Technologies.

The variation in responses from the panel, with respect to both format and resolution, increased the requirement for the author/facilitator to post-process responses. While templates were provided for responses, their format were largely open to interpretation by the expert panel. This was exemplified in round 2 where the panel were asked to identify interactions between variable pairs, potentially a time-consuming task for the panel members. This is reflected by the overall feedback supplied by Expert G that "it [the Delphi study] took more time than anticipated to complete". A subsequent application of the Delphi Method this FMC method or more generally, should therefore include more explicit expectations for the format of responses to be provided for each round, with a view to ensuring both quality of the output, and ease of participation. Whilst not applicable to questions where free text is required, and indeed desired, for quantitative questions such as identification of interacting variable pairs, or assignment of relative significance, further clarification of the expected response format would decrease any judgement required by the facilitator, while possibly decreasing the time and effort required by the expert panel. However, to best elicit the qualitative expert knowledge from the study, sufficient opportunities to provide justifications, comments, and evidence alongside any responses should be retained.

In this case study, some members of the panel demonstrated a reduced willingness to provide qualitative justifications and evidence alongside their judgements, to the detriment of the effectiveness of the overall study. This may demonstrate either an uncertainty associated with their expert judgement, an unwillingness to participate fully, or that they were more time bound in their response. This had implications in both the groups' assessment of the significance of the variables, and the facilitators understanding motivations behind changes of judgement.

Furthermore, despite the expert panel being selected in part due to their perceived motivation and availability to participate, two experts elected to opt-out, while others failed to contribute either a response, or an apology for multiple rounds. Despite being identified by the stakeholders as somewhere with prior relevant experience, Expert D elected to opt out of the study in its entirety. This may reflect the expert selection criteria, or their own perception that their current role outside of active product development meant that they could not impart any relevant knowledge. This suggestion is supported by the fact that Expert H, who was a member of the same functional group at the time, only completed one round of the study before then failing to contribute further. In future iterations of this method, additional emphasis should be made that knowledge of a failure mode on previous products is as important as for current products.

The other panel member to opt-out, Expert I, did so after the first round, taking issue with the responses of some of their peers, stating that the responses were either ‘too vague or too specific’. This was discussed in person with Expert I, reminding them of the iterative nature of the Delphi study, and the opportunity it would present to better characterise the failure mode with their peers. Unfortunately, despite this, the expert elected to take no further part in the study. This may be a reflection on the appropriateness of the expert selection criteria, or of the nature of the personality of the expert in question.

Two experts, F and L, did not respond to any round, nor did they choose to opt out. One-on-one communication with each expert resulted in repeated apologies and verbal commitments to participate, suggesting workloads as the cause for delay, but time constraints prevented them from participating in the study at all. This may suggest that additional care would need to be payed to the ‘availability’ element of the expert selection criteria in the future to best ensure panel members would be able to participate.

Expert E represented the only member of the Analysis & Simulation group, and only participated in the second round after providing an apology in the first round owing to workload. Given that previous investigations into this failure mode involved significant use of CAE tools, this may represent an opportunity for significant knowledge to be missed from the Delphi study. However, as CAE studies are heavily integrated into the product development process at Delphi Technologies, the impact of this reduced participation was mitigated, supported by other expert panel members submitting evidence and justifications in the form of CAE results and reports.

Despite a clear and justified request for anonymity, several members of the expert panel discussed the study with their colleagues and peers in open plan office spaces. This in turn led to several panel members becoming aware of each other. Furthermore, other panel members commented to the author that they could identify some of their peers through the justifications and evidence they provided. These instances highlight a potential limitation of an internal, homogenous panel with respect to the anonymity that is a key element of the Delphi Method and one of the firmest considerations for its application to this case study.

One expert rejected the ‘pigeon holing’ associated with the ranking of the variables identified by their peers, which they considered as treading old ground. This was met with the reply that this was an opportunity for them to challenge the beliefs of their peers anonymously, after which he continued his contribution, demonstrating the benefits of the anonymity afforded by the Delphi Method.

This case study provided evidence to support the concern that technical experts engaged in customer focused NPD programmes are perhaps focused too strongly on emergent problems, to the detriment of reflection. The expert panel consistently showed an unwillingness to reflect on the results of previous rounds, such that the design of the study was changed to effectively 'force' more reflection on the relative significance of the variables. This is evident in the fact that only one expert referenced their judgement on the variables back to the group led definition. This lack of reflection and iteration led to many of the variables defined as either being 'too vague' or 'too specific' as identified by one panel member, which then was further evident in the acknowledgement by some panel members that several of the variables where no consensus could be met were open to interpretation.

The study also demonstrated some of the difficulties with the facilitator being too 'close', both to the panel of experts and to the problem. The professional relationships between the facilitator and some members of the panel became strained, leading to some experts avoiding personal contact after they were unable to complete a round, which would not be a consideration if the facilitator was unknown to the panel. The corollary to this is that the facilitator's professional and personal relationships with the panel may have assisted in eliciting candid responses. The facilitator's closeness to the problem was acknowledged when analysing the qualitative data, where the significance of their own expert biases had to be recognised. As such, minimal 'post–processing' was applied to the textual results to limit any influence of their own judgements.

After the study had completed, the panel were asked to provide feedback on the study itself. As previously mentioned, Expert G highlighted that the study proved more time consuming than they had anticipated. Expert G also commented that "It was very interesting to hear others opinions and they did influence my own", while Expert A commented that "[the Delphi study was] an effective method overall". Reflecting on the diversity in the opinions around the failure mode and the nature of some experts within the business, Expert B commented that "a workshop [as an alternative method] would definitely not have been effective: it would have been a blood bath".

In an industrial context where it can be difficult to achieve and maintain participation in FMEA and 8D groups, it is suggested that the Delphi Method could offer a resource sensitive approach to eliciting expert judgement, supported by the expert community within Delphi Technologies. However, examination of this application, alongside the feedback of some panel members, have identified a number of potential areas for iterative improvement with regards to the design and implementation of the Delphi Method in Delphi Technologies.

From to perspective of the wider research objectives, this application of the Delphi Method has been successful in generating a definition of the failure mode, alongside a description of the variables that influence it and their perceived significance, based on the aggregated knowledge of a cross section of experts from within the business.

More generally, the purpose of the EE element of this FMC method is to kick-off or restart the discussion going within an expert community, allowing discovery and combination of tacit knowledge. In this case study, the Delphi Method has been presented as an effective method for doing so but may not represent the most appropriate EE method in all applications.

9.2.2 Reflections on the application of Causal Loop Diagrams as a system modelling tool within the overall methodology

The system models of the failure mode have resulted in effective boundary objects that are suitable for translation into parameterised models for further validation through experimental investigation. Due to the high number of contributing variables and their interactions, associated with the failure mode, the complexity of the figure increased significantly through the addition process. This rich detail did limit the suitability of the model as a boundary object for use with some audiences, proving intimidating in the amount of information being presented, often met with bemusement and derision, questioning both the legitimacy of the model, “Is that real?”, and the suitability of the model, “Can you follow that?”.

The original iteration of system model of the failure mode only included those variables present in the qualitative feedback from the expert panel, representing what they were ‘talking about’, rather than the comprehensive list of variables identified in the formal element of the study. That original model excluded 6 variables all identified as being of high significance to NCV seat wear, so their omission from the descriptive feedback provided by the panel may provide additional insight to how they are perceived. The four design variables, and the one manufacturing variable omitted from the original model, all influence the ‘As manufactured seat condition’, a variable that was not itself identified as being of high significance to NCV seat wear. This suggests that while the panel identify the design and manufacture variables that influence that as manufactured condition, there perhaps exists sufficient control of tolerances and demonstrated capability in manufacturing processes as to reduce the perceived risk associated for NCV seat wear. An alternative suggestion would be a perception within the product engineering functions that their design and manufacturing techniques are adequate, and that the liability for the failure mode is associated with usage, and therefore out of their control.

The omitted usage variable, ‘Hard particle debris contamination’, may be deemed controlled through a combination of the warranted fuel specifications agreed with Delphi Technologies’ customers, or through the performance of filtration systems on Euro VI engines. Furthermore, hard particle debris contamination is typically associated by the panel with seat wear resulting from flow erosion, outside of the agreed root definition that resulted from the Delphi study in §6.4.2.

The systems model of the soft system causality that has prevented Delphi Technologies from fully characterising this failure mode in the past validates several aims of this research, while also proving an effective boundary object for engaging stakeholders. The model identifies causal relationships where the core competencies of the business could be improved, presenting a description of the perceived problem, and allowing the desired potential end-state of the system to be communicated, facilitating engagement with stakeholders with regards to identifying appropriate interventions.

In a similar manner to the application of the Delphi Method, this implementation of Causal Loop Diagrams has highlighted both benefits and limitations associated with the author/facilitator being close to the problem. Firstly, in performing the qualitative data analysis required to construct the causal fragments for each model, the author/facilitator performed minimal post-processing of the results to reduce the influence of any of their own expert bias, with the potential to reduce the effectiveness of the results of that analysis. However, for the hard system model, the author/facilitator's closeness allowed relationships between variables governed by first principles to be readily identified, negating the requirement for their own expert judgement to be applied to bridge gaps in the model. For the model of the soft system causality, where no such first principle relationships exist, and the number of responses received was decreased, the author/facilitator's own expert judgement was required to bridge the gap between causal fragments such as to create a system model, potentially introducing their own expert bias.

In §4.5.1, FTA was presented as tool for RCA typically employed in industry. The figure that results from FTA, or similar hierarchical failure models, can be compared to the CLD generated in this case study as they both attempt to capture the different variables that can influence a failure mode and attempt to describe the mechanisms through which it influences the failure mode. Figure 126 shows an example of a diagram resulting from a hierarchical failure modelling process, and while the figure itself is deliberately illegible, the highlighted nodes represent the same factor repeated a total of 9 times across 4 different high-level branches of the Fault Tree. In comparison to the CLDs presented in Chapter 6, an FTA diagram, or similar, does not provide a sense of variable interactions, or adequately describe how effects can propagate through a complex system. It is suggested that CLDs offer an effective alternative to FTA diagrams for capturing that dynamic complexity as part of systematic problem structuring.

Figure 126: Example FTA diagram with repeated factor highlighted[4]

9.2.3 Reflections on combining the Delphi Method and Causal Loop Diagrams

The Delphi Method has been demonstrated as an effective means of generating the elements of a useful system model of a failure mode in the form of a CLD. In identifying variables that influence the failure mode, along with their interactions, and perceived significances, the Delphi Method can serve as an appropriate method for combining tacit knowledge into a system model. Future iterations of this methodology could possibly benefit for further integration of the two component methods to both facilitate the implementation of the methodology, and to improve the quality of the results.

To better engage an expert panel in the qualitative feedback, particularly when using the Delphi Method to construct systems models, causal fragments should be included as part of the aggregated results for appropriate questions. Through visualisation of the feedback in the form of system model elements, the panel may perhaps be more engaged with the explorative element of the Delphi Method. By doing so, the quality and quantity of qualitative feedback should then improve, and when using the Delphi Method, the best way to deal with incomplete data is to prevent it in the first place (Geist, 2010).

When using the Delphi Method to construct systems models, either of hard or soft systems, a final group model building workshop may be beneficial in refining the model. Any gaps in the model that remain after the completion of the Delphi study would typical require the facilitators own expert knowledge, through either identification of the appropriate first principle relationships that govern elements of the system, or their own expert knowledge. This model building workshop should be

[4] This figure is deliberately illegible as it is taken directly from a real Delphi Technologies application

completed with either the expert panel, or key project stakeholders, and will result in a refined model that best represents the system, while being a purposeful boundary object.

More generally, this case study has suggested that Causal Loop Diagrams represent an effective means for codifying the results of an EE programme into a useful boundary object for the NPD process, representing a visually elegant way of modelling the rich knowledge associated with a failure mode and effectively capturing and communicating the complexity of a problem.

9.3 Reflections on Experimental Design as an explorative element of a systematic approach to failure mode characterisation

This section will reflect on the empirical investigation through experimental design methodology discussed in Chapter 7.

9.3.1 Reflections on experimental design

The Delphi Study and subsequent modelling of the system proved an effective precursor for experimental design. In combining expert judgement from around the business, variables and their interactions were made available, along with perceived levels of significance. Furthermore, the mechanisms that result in the failure mode were established, together with how the failure mode manifests in usage were determined. In doing so, both the potential design factors and response variables has already been established.

To determine design factors, and their levels, that were both appropriate and robust, some pre-experimental investigations were required, but it is recognised that the experimentation is iterative, as indeed is the systematic approach presented in this thesis. Most significantly, in the experimental design process, a significant gap in the knowledge of fuel properties, with respect to both usage conditions and degradation, was identified within the Product Engineering function of Delphi Technologies.

The factorial design methodology selected for use in this case study provided a resource effective means for exploring both factor effects and interactions, with the possibility of also exploring the linearity of effects. However, it is recognised that for further applications of this approach, alternative Experimental Designs may prove to be more suitable depending on the number of factors to be investigated, and the available resources in industrial environments.

Replication is one of the main principles of Experimental Design methodology, allowing full interpretation of the effects of experimental error. While resource availability prevented any full replication of tests, the experimental design did include forms of pseudo-replication that served to increase the confidence associated with any conclusions. Firstly, in using automated test plans, with both design factors and covariates controlled dynamically over the duration of the test, experimental error associated with test conditions was minimised. Additionally, the random samples chosen for the tests were all built using the same highly controlled assembly processes, reducing any associated experimental error associated. Furthermore, in using 12 samples for each treatment combination, in turn fitted to two parallel FIS, additional elements of pseudo-replication serve to provide a quantification of experimental error. Finally, through design projection, experimental error can be approximated through the removal of the least significant effects.

The decision to consider only usage variables, and thus considering design and manufacturing variables as fixed within normal production variation, allowed for a resource sensitive approach to test design, but necessitates further investigation. However, in doing so, the knowledge gained from the experimental programme will serve as the foundations of a controlled and repeatable accelerated test that could be used to evaluate future robust design solutions.

9.3.2 Discussion of experimental results

The empirical testing programme considered 3 different usage variables identified by the Expert Panel as being of high significance to NCV seat wear. With statistical confidence, only one of those variables, *System Operating Pressure*, is observed to result in a significant response in physical NCV wear and injector performance drift, while *Fuel Type* and *Number of Injections* are insignificant. However, a secondary statistically significant effect can be observed on the timing of the injector as a result of varying the Fuel Type factor.

Significant location effects

Increasing the System Operating Pressure has been demonstrated to result in an increase in both the Trimmed Integral of Drift injector performance metric, and the physical seat wear of the NCV. The system operating pressure influences several factors, which can in turn influence NCV seat wear. Figure 127 presents an extract of the system model, demonstrating how variations in system operating pressure propagate through the system, potentially influencing NCV seat wear.

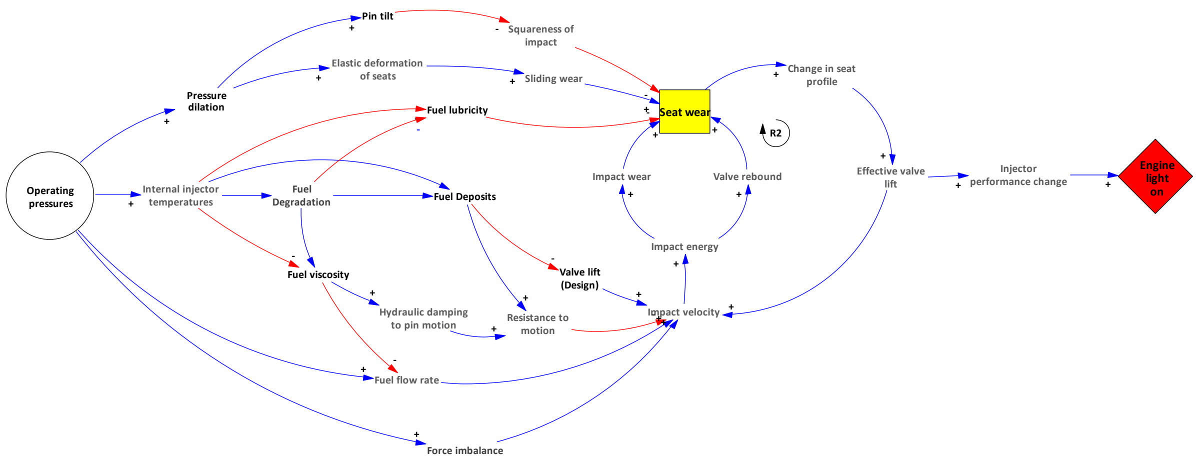


Figure 127: Extract of system model demonstrating causality associated with operating pressure

As the operating pressure can influence both impact and sliding wear, additional evidence may be required to fully understand the effect on the failure mode. As the impact wear is only associated with the actuation of the valve, the control samples can provide some insight into the relationship between operating pressure and seat wear associated with sliding wear. Figure 128 visualises the location of each control sample at 1000hours relative to the sample mean (with 95% confidence intervals) for both the TI drift metric, and the NCV seat wear metric. As can be seen, for the tests conducted at higher operating pressure, tests 2 & 4, both a performance change and a measured NCV seat wear can be observed, while the tests lower operating pressures show little-to-no change from as new condition. Furthermore, the difference level observed in both metrics at higher pressures when compared to lower pressures, is similar to the difference observed in the sample means for the same tests as represented in Equation 12 and Equation 14.

$$\left(\overline{TI@2500} - \overline{TI@1800}\right)_{Test\ Samples} \approx \left(\overline{TI@2500} - \overline{TI@1800}\right)_{Control\ samples} \tag{12}$$

$$\left(\overline{NCV@2500} - \overline{NCV@1800}\right)_{Test\ Samples} \approx \left(\overline{NCV@2500} - \overline{NCV@1800}\right)_{Control\ samples} \tag{13}$$

Equation 12 could be rearranged such as to suggest that the difference in magnitude of TI drift for the test samples is approximately equal to the difference in magnitude observed in the control samples, as shown in Equation 14.

$$\begin{aligned}\left(\overline{TI@2500}\right)_{Test\ Samples} - \left(\overline{TI1800}\right)_{Test\ Samples} \approx \left(\overline{TI@2500}\right)_{Control\ Samples} \\ -\left(\overline{TI1800}\right)_{Control\ Samples}\end{aligned} \tag{14}$$

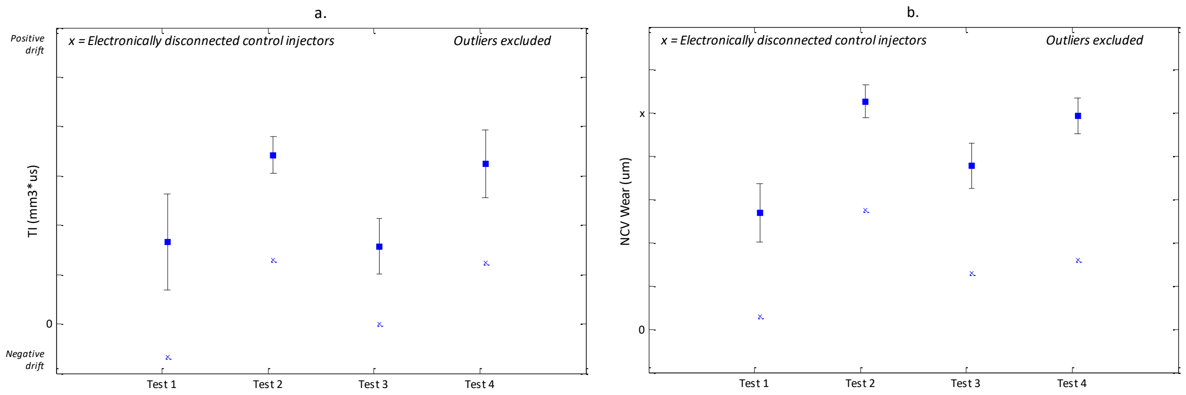


Figure 128: Location of control samples against sample means: a.) TI Drift, b.) NCV seat wear

The control samples were only actuated during the 3 performance tests, with no injection events taking place during the 1000-hour test. Since the number of actuations associated with the performance test (~45,000) represents only approximately 0.03% of the actuations for the test itself (144,000,000 for the test with 2 injections per cycle), it is suggested that any wear associated with the performance tests would be insignificant compared to the wear associated with the test itself. It is therefore proposed that the difference in location can be observed between both the performance, and the physical seat wear itself associated with control samples associated with different pressures, is because of a component of wear not associated with actuation of the valve, and impact wear on the valve seats.

Sliding wear of the control valve can be associated with elastic deformation of the pin, both as a result of valve closing in the form of compound wear, and through the varying pressure difference, and resulting force balance acting on the valve in static conditions. While the control samples would not be exposed to any sliding wear that may result from impact of the valve seat, being exposed to pressurised fuel, they were still exposed to variations in pressure, and as such, it is proposed that the wear, and performance change, observed in the control samples may be a result of sliding wear associated with pressure dilation, as described in Equation 15.

$$(TI, NCV)_{Control\ Samples} \propto Sliding\ wear\ from\ pressure\ dilation \qquad (15)$$

The increased levels of wear associated with the test samples is therefore proposed to be a result of the contribution to wear associated with valve impact, either in the form of plastic deformation, or through elastic deformation and radial sliding of the pin seat in closure, as described in Equation 16.

$$(TI, NCV)_{Test\ Samples} - (TI, NCV)_{Control\ Samples} \propto Impact\ wear \qquad (16)$$

As Equation 14 demonstrated, the difference in wear, and performance, between the test samples associated with each pressure was approximately equal to the change in wear and performance observed between the control samples associated with each pressure. As such, it is proposed that the NCV bottom seat wear, and the element of injector performance drift associated with it, of an injector can be described as a combination of a sliding wear dependant on the operating pressure of the system, and a component of impact wear associated with the operation of the system, as described in Equation 17.

$$NCV\ Bottom\ Seat\ Wear\ \propto Sliding\ Wear\ (Pressure) + Impact\ Wear\ (Operation) \tag{17}$$

In §6.4.2, the expert panel agreed a definition of NCV seat wear that only included a sliding wear mechanism. The results of the empirical study support that definition in part, identifying the component of the failure mechanism that varies with usage condition as sliding wear. However, the empirical study suggests that there is a component of the failure mode related to impact wear, which does not appear to vary with usage conditions.

Further empirical testing could provide additional insight into the nature of the wear mechanism and its components. The empirical testing associated with this study was limited to one control sample per treatment combination, and as such, provide a low statistical confidence for any associated conclusions. Further experimentation, with an increased proportion of electronically disconnected control injectors could provide additional insight into both the difference observed between test and control samples, and the effect of usage variables on the response of control samples. Such experimentation should be focused on the system operating pressure as the only of the usage variables to influence sliding wear from dilation and could serve as a suitable test for alternative design variables to characterise the effects of seat form and pin stiffness.

Additional empirical investigation could also further the understanding of the component of wear associated with valve actuation. To isolate any effects of sliding wear associated with pressure dilation, a series of 'dry' experiments could be performed, using a controlled, representative pin impact force to examine the component of wear associated with pin impact, while a cyclical mechanical loading of the valve in the closed position could investigate any wear mechanisms associated with the valve elastically deforming on closure. Both investigations would need to be supported with analytical investigations to determine the mechanical loads required to provide representative forces and energies associated with valve actuation.

The fuel type variable is shown to have a significant location effect on injector timing, as expressed through the minimum drive pulse variable. Since the two fuels used are shown to exhibit differing levels of chemical degradation, and thus IDID formation propensity, through recirculation, it is suggested that the timing retardation observed on the tests using WWLTF is because of deposit formation. Further investigation into the level of any deposits observed in the NCV assembly could provide additional confidence in this conclusion.

Significant dispersion effects

The empirical testing programme identified the number of injections as being a significant effect on the dispersion of both drift variables, where an increasing number of injections resulted in a decreased level of dispersion. Since no such significant effect is observed in the physical wear of the NCV or PG, it is suggested that this is the result of a separate wear mechanism influencing injector drift but not NCV seat wear as quantified through measurement traces.

One such potential wear mechanism is that associated with hard particle debris damage. As discussed in §3.7.3, debris damage occurs either through free particles striking the seat as a result of fuel flow, or through particles being trapped between the seats during valve closure. Both such mechanisms require the seat to be open and are thus linked with the actuation of each injection. Increasing the number of injections, will then increase the probability of debris damage for a valve seat for a given test duration.

It is therefore proposed that the significant effect observed in a decreasing dispersion of drift metrics associated with increasing number of injections, is because of increased probability of debris damage. When comparing the samples tested with 6 injections per cycle with those tested with 2 injections per cycle, it is suggested that there will be a higher mean level of debris damage observed, but also a lower level of dispersion, as the probability of any seat exhibiting debris damage is increased through higher numbers of valve actuations.

Initial investigation, through quantifying the number of significant debris damage sites on NCV pins, has suggested this hypothesis may be valid, but with a low level of statistical confidence. Figure 129 compares the distributions of the two sample populations, while Table 30 provides the descriptive statistics for each. As can be seen, when an outlier is excluded, using the 1.5 times interquartile range method, the mean for the 2 injections per cycle samples is lower than the mean for the 6 injections per cycle samples, while the standard deviation for the 2 injections per cycle samples is higher.

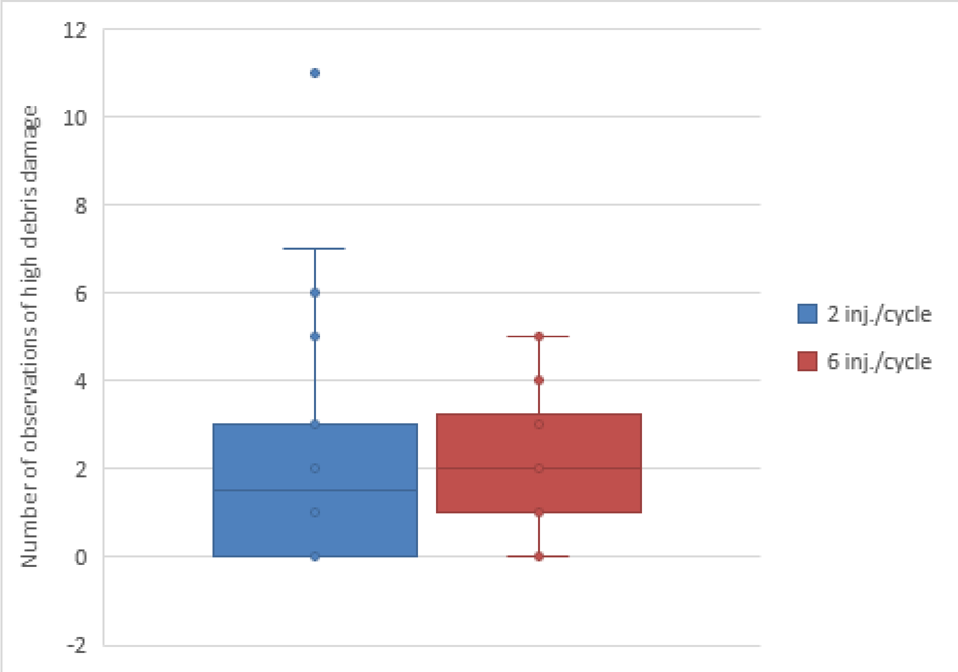


Figure 129: Comparison of distribution of debris damage for different numbers of injections/cycle

Debris Damage Observations	2 injections/cycle	6 injections/cycle
Count (exc. Control samples)	22	22
Mean	2.2	2.1
Std. Dev.	2.8	1.7
Minimum	0	0
Maximum	11	5
Range	11	5
Median	1.5	2
Interquartile (25%) - IQ1	0	1
Interquartile (75%) - IQ3	3.0	3.5
Interquartile range (IQR)	3.0	2.5
IQRx1.5	4.5	3.75
IQ3 + 1.5xIQR	7.5	7.25
Mean (exc. Outliers)	1.8	2.1
Std. Dev. (exc. Outliers)	2.1	1.7

Table 30: Comparative statistics of distribution of debris damage for different numbers of injections/cycle

Additional investigations, quantifying the level of material removal associated with debris damage for each sample as a more robust means of assessing debris damage, and ultimately conducting additional testing with different levels of debris concentration, could provide additional confidence in this hypothesis.

Insignificant location effects

Varying the number of injections had no significant effect on the location response of either the drift, or physical wear metrics. As discussed previously in this section, analysis of the results of the control samples suggest that the element of seat wear, and injector drift, that varies with usage variables may be associated with sliding wear resulting from pressure dilation rather than valve actuation. As such, that wear component would not further vary with number of injections. Of the remaining component of wear associated with valve actuation, initial interrogation of the data suggests an approximately equal amount of wear and drift associated with impact for samples with both 2 and 6 injections per cycle. This suggests either a plateauing of that component of wear, or that the low logic pilot and post injections associated with the 6 injections per cycle tests, result in different impact conditions such that the main contribution of wear is associated with the long logic main injections, common to both tests. Additional empirical investigations, firstly observing the relationship between the number of injections and the actuation component of wear, and secondly, observing the relationship between injection logic and the actuation component of wear, may provide additional insight into the significance of any relationship between the injection strategy used and NCV seat wear.

Varying the fuel type, primarily considering lubricity, but also varying viscosity, had no significant effect on either the drift, or physical seat wear metrics. With respect to the fuel lubricity, this lack of any effect could be because of recirculation and degradation negating any significant differences between the fuels, a result of internal injector temperatures influencing the effective fuel lubricity at the valve seats, or as a true result of the fuel lubricity having no effect on the wear mechanisms present.

As discussed in §7.3.4, while WWLTF does represent at lower lubricity fuel than ISO4113 while considering the effects of fuel degradation, the difference between the two levels is not as significant as desired. Further empirical investigation should therefore be conducted with fuels with an increased separation in lubricities, through either use of non-recirculated fuels, or through fuel additives intended to either reduce fuel degradation or reduce fuel lubricity. However, additional care must be placed on limiting the propensity for IDID formation, the effects of which can confound any lubricity related effects.

As discussed in §3.5 the lubricity of fuels can vary with temperature, with low lubricity fuels demonstrating worsening lubricity up to ~100°C, before then improving back to a level equivalent to that observed at lower temperatures. However, that investigation was conducted using an HFRR test facility in which the fuel is open to the atmosphere. The authors noted that the inflection observed around 100°C was unexpected and may have been as a result of any water content or products evaporating out of the fuel (Lacey et al, 2001). In a closed system, such as the internals of a DFI, any water or other products would be unable to evaporate from the fuel, so the observations from experimentation on HFRR rigs may be unrepresentative. Furthermore, internal injector temperatures at 2500bar, would be such that the fuel would be significantly beyond the point of inflection observed in HFRR testing. As such, uncertainty exists around the relationship between fuel lubricity and temperature in closed systems such as the DFI. Additional characterisation of the behaviour of fuels is therefore required to fully understand the potential relationship between fuel lubricity and wear.

A further source of uncertainty around the potential effect of fuel lubricity on NCV seat wear concerns the ability of fuels to form tribofilms on DLC surfaces, and indeed, the NCV bottom seat interface may instead be lubricated by the coatings alone. Numerous studies have been conducted using both lubricant oils, and fuel oils, on coated sliding surfaces to observe the effects on wear and coefficients of friction. However, conflicting information exists concerning the ability of mineral oil to form an effective tribofilm on the specific type of coating[5] used on the NCV bottom seat interface. Lubricant oils typically include additives specifically designed to promote tribofilm adhesion on a range of coated surfaces, but such additives are not employed in fuel oils. Since hydraulic test rigs do not require fuel combustion, it may prove possible to conduct further empirical testing using fuels with such additives to assess the effects, if any, on bottom seat wear.

The expert panel who participated in the Delphi Study, failed to identify any such uncertainty around the lubrication of the NCV bottom seat, reaching a consensus on it being of high significance to NCV seat wear. This fact may firstly reflect a deficiency of the composition of the expert panel itself, suggesting the inclusion of experts with more specific fuel and tribology knowledge, rather than product knowledge, may have proved complimentary, to the benefit of both the efficacy of the study, and to the knowledge of the expert panel.

Metrics for Seat Wear and Drift

One objective of the empirical investigation was to identify injector drift metric(s) that could represent a means for inferring progression of the failure mode without physical disassembly and measurement. Figure 130 shows the directionality of the location and dispersion effect for each combination of usage variable and metric, with effects demonstrated to be statistically significant highlighted. By comparing the two-performance metrics to NCV seat wear, conclusions can be made about their relative suitability. With regards to the UM drift metric, the directionality of location responses can be seen to be the same as that observed in the NCV seat wear metric, but the significant effect observed in the

[5] While the type of coating used cannot be disclosed, a review of the literature will highlight the conflicting findings regarding different types coatings, including those doped with other elements.

rail pressure variable for seat wear was not correctly identified. In comparison, the TI drift metric incorrectly identifies the directionality of the response associated with fuel type but does identify the significant effect associated with rail pressure. Both performance metrics correctly identify the directionality of dispersion effects. The MDP metric is shown to exhibit opposite directionality than NCV seat wear for each usage variable.

Location	MDP	UM	TI	NCV	PG
Fuel	**↑**	↓	↑	↓	↑
Rail pressure	↓	↑	**↑**	**↑**	↑
No. of inj.	↓	↑	↑	↑	↑

Dispersion	MDP	UM	TI	NCV	PG
Fuel	↑	↑	↑	↑	↑
Rail pressure	↓	↓	↓	↓	↑
No. of inj.	↓	**↓**	**↓**	↓	↑

Legend
↑ Raising level results in increased response
↓ Raising level results in decreased response
Significant response

Figure 130: Comparison of effect directionality and significance for different metrics

When viewed in comparison to NCV seat wear, and in the context of rail pressure being the only significant variable, the TI drift is the most effective of the injector drift metrics, and thus could be used as an effective means for inferring NCV seat wear over time in Accelerate Degradation Tests. However, the metric itself is more abstract than the existing UM drift metric, so the UM drift metric would likely be used in parallel for a transition period. The MDP metric can be shown to have little-to-no relation to NCV seat wear but does provide a useful means for inferring fuel related effects.

9.3.3 Reflecting on the assumptions associated with the experimental element

In §7.2.2 four assumptions were presented as part of the design of the experimental element of this case study.

The first assumption was that NCV seat wear can be accelerated on a hydraulic test rig using recirculated fuel. The results discussed in this Chapter demonstrate that this assumption was valid, and that differing levels of NCV seat could be observed for the same run time on a hydraulic test rig.

The second assumption was that NCV seat wear can be accelerated through variations in usage variables while not being confounded by other failure modes. As discussed in this Chapter, this assumption has been demonstrated to be valid in part, in that the location of the response in NCV seat wear can be shown to vary with rail pressure, but the dispersion in responses is yet unattributed.

The third assumption was that the expert elicitation study had identified the usage variables that have the most significance on NCV seat wear. As discussed in this Chapter, the expert panel correctly identified that Rail Pressure had a significant location effect on NCV seat wear and the performance metrics, but no other usage variables that they identified have been demonstrated as having a significant location effect. The number of injections variable was however shown to have a significant dispersion effect on the performance change metrics, but not the physical wear metrics. As such this assumption has been demonstrated to be valid in part.

The final assumption was the NCV seat wear could be correlated with Injector performance change. As discussed in this Chapter, while the prior metrics for drift were shown to be unsuitable, the TI drift metric has been shown to be an effective inference of NCV seat wear, while a correlation between the two metrics was shown in §7.12. As such, this assumption has been demonstrated to be valid.

9.4 Discussion of regression modelling element

9.4.1 Assumptions associated with the generalised model for TI drift & potential mitigations

The first assumption associated with this model is that ***the only significant usage factor associated with TI drift is Rail Pressure***. While the DoE testing completed to date suggests that neither fuel type, nor number of injections are statistically significant factors, additional testing, and the inclusion of other usage variables, would further improve the statistical confidence associated with this assumption.

The second assumption is that ***the relationship between TI drift and rail pressure is non-linear, and that the Michaelis-Menten regression model can be used to describe it***. The single pseudo-centre point suggested with a >85% confidence that curvature was present in the rail pressure effect for TI drift, but additional test points within, and outside of, the experimental boundary will provide further confidence in this assumption. The same testing will then improve the confidence associated with the regression model, using either the Michaelis-Menten model, or another model as appropriate.

The third assumption is that ***TI drift is non-linear with time, tending towards an asymptote, and that the Michaelis-Menten model can be used to describe it***. Additional performance inspection intervals, for all values of hours, will improve the confidence in the assumption of non-linearity and the suitability of the regression model used to describe it, being either the Michaelis-Menten model, or another model as appropriate.

The fourth assumption is that ***the TI drift observed at 1000 hours for different rail pressures is proportional to the asymptotic value, defined in the Regression model by Theta 1, and that relationship can be described using linear regression***. A linear regression model was fitted to the available empirical test data, but further development of the regression models for multiple pressures will potentially improve the confidence associated with this assumption.

The fifth assumption is that ***the shape of the regression curve at low hours, defined in the regression model as Theta 2, is constant, meaning that the same percentage of the asymptotic value is always reached after the same number of hours, regardless of pressure***. Additional performance inspections at all rail pressures at low numbers of hours will improve the understanding of the relationship between pressure and the shape of the regression curve, described through the Michaelis-Menten model or otherwise. If validated, this relationship will facilitate reduced duration testing.

9.4.2 Limitations associated with the generalised model of TI Drift

The first limitation is that ***the model cannot predict TI drift for rail pressures lower than 1800bar***. This is because of the regression model used to describe the relationship between TI drift and rail pressure, and additional testing at pressures below 1800bar would allow the model to be redefined such as to provide any necessary predictions.

The second limitation is that ***extrapolation would be required to predict TI drift for any pressures greater than 2500bar, with confidence levels for any such predictions unknown***. Additional testing outside of the existing model boundary will allow confidence bounds to be placed on such predictions, while improving the suitability of the generalised model for all rail pressures.

The third limitation is that ***the model underestimates the TI drift at higher hours***. Additional performance inspections for all values of pressure and temperature will improve the adequacy of the model for prediction TI drift across the full range of pressure and hours run.

The fourth limitation is ***that the model does not predict the dispersion associated with TI drift, only its location, and that the dispersion will be such that the highest drift injectors for a lower pressure test may have similar levels of TI drift as the lowest drift injectors from a higher-pressure test***. Additional development of the model could provide a prediction of the dispersion in TI drift, while additional understanding of any other design and usage variables that influence that dispersion could improve the overall understanding of the failure mode.

9.4.3 Comparison of applications

In §8.4.1, different application duty cycles were compared using the generalised model of drift. It was shown that each duty cycle would result in different progressions of fuelling change as vehicle km, or hours, were accumulated as a function of their differing rail pressure residencies. In doing so, it was shown that the most severe duty cycle can result in an acceleration factor of ~10x over the least severe. While the predictions associated with the different drive types would need to be validated with field returns, with emphasis on the knowledge of the application and its typical drive environment, this analysis uses the generalised model of TI drift to demonstrate the magnitude of the influence exerted by the application duty cycle on the rate of progression of the failure mode.

Given that all were on-highway applications and would therefore be granted equal FIS warranty conditions based on vehicle kilometres, this then highlights a potential shortfall in the current warranty statement as not all applications are equal. In the future, with increasing sophistication of on-board electronics, it may be desirable to instead base warranty conditions on specific application usage, rather than a generic vehicle distance, or more simply, base FIS warranty on vehicle classification and likely duty cycles.

9.4.4 Why does the previous mental model suggest engine is more severe than rig?

In §6.2 a working mental model for drift progression on both rig and engine tests was presented, in which, the drift, more specifically the component of which is associated with NCV seat wear, associated with testing on hydraulic rigs was suggested as plateauing at a lower level than on engine-based tests. This empirical investigation has yielded new knowledge on the factors that influence drift, so this problem can be reviewed, and several hypotheses can be suggested to explain that mental model further.

Firstly, the existing metric for injector drift has been demonstrated to not be the best metric for differentiating between sample populations, or for inferring NCV seat wear. The mental model for drift progression used untrimmed maximal ballistic drift, a metric that is inherently cofounded with any timing changes, and which can result in high levels of drift being observed, but only associated with isolated points on the gain curve. The trimmed integral of drift metric has been suggested as a more appropriate means for describing changes to the shape of the gain curve without being confounded by any timing changes. TI drift has also been demonstrated as a more appropriate metric for observing the differences between sample populations, and best infers NCV seat wear. As such, it is suggested that the existing mental model for drift was developed using a sub-optimal metric.

Secondly, this research has demonstrated that testing with high time residencies at higher pressures results in an increased level of drift and seat wear. A significant majority of the testing on hydraulic rigs at Delphi Technologies, particularly longer hour tests, is completed on cyclic test profiles, in order to accelerate the system as a whole. This majority forms the basis for the mental model of drift associated with rig testing, but this research has demonstrated that it results in a decelerated progression of drift when compared to steady state testing at high pressures only. Customer testing on engines encompasses a broad spectrum of test profiles, including combustion development work, and emissions optimisation, that may include significant time residencies at steady state conditions, and in a similar manner to rig testing, can accumulate hours faster than a vehicle. As such, it is suggested that a portion of engine tests, conducted with high time residencies at high rail pressures, do represent an effective acceleration over cyclic rig testing, which coupled with the heightened sensitivity associated with problems experienced by customers, influences the mental model for drift.

Thirdly, while this research did not conclude that fuel type as had significant effect on drift, hydraulic rig testing uses recirculating fuel, that improves the effective lubricity of fuels. Fuel recirculation has been demonstrated to result in an improvement of the lubricity of more aggressive fuels through fuel degradation. As such, for a hydraulic test rig, the lubricity of the fuel improves over time, so only a small proportion of the testing is conducted with a low lubricity fuel. Unlike hydraulic rig testing, only a very small proportion of the fuel, associated with system leakage or rail pressure control, used for an engine test recirculates back to the fuel tank for further usage, negating the effects of any fuel degradation. As such, it is suggested that engine tests can be conducted with a fuel of lower effective lubricity than hydraulic rig testing.

9.4.5 What may be the result of different technological trajectories?

Using the equation of TI drift in terms of rail pressure presented in in §7.11, the expected values of TI drift after 1000 hours of run time can be estimated, as shown in Table 31, where the italicised values are those outside of the bounds of the model.

Rail Pressure (bar)	**TI Drift (mm3*us)**	**Percentage of TI drift @ 2500bar**
1800	314	45.8%
1900	491	71.6%
2000	569	83.0%
2100	613	89.4%
2200	641	93.5%
2300	660	96.3%
2400	674	98.4%
2500	685	100.0%
2600	*694*	*101.3%*
2700	*701*	*102.3%*
2800	*707*	*103.2%*
2900	*712*	*103.9%*
3000	*716*	*104.5%*

Table 31: Relative predictions of TI drift at different pressures

Considering the extrapolated values, and the asymptotic behaviour of the relationship between TI drift and rail pressure within the boundary of the model, it can be suggested that increases of rail pressure above the current level of 2500bar, would potentially not result in large increases of fuelling drift observed. However, testing outside of the existing experimental boundary would be required to confirm that behaviour. Alternatively, the benefits in lowering the rail pressures in application can be clearly observed, with a >50% reduction in TI drift demonstrated within the current experimental boundary.

One of the aims of FMC is to identify and validate potential robust design solutions. The model generated in this case study suggested that with design optimisation, it may be possible to either increase rail pressure further while not resulting in an increase in fuelling drift over time or utilise lower rail pressures to enable a significant improvement in product stability over life.

9.5 Discussing the gaps that exists after the empirical investigation

Arguably the most significant knowledge gap that exists after empirical testing is the influence of the design and manufacture of the NCV. This was a result of a deliberate decision when determining the boundary of this investigation but was one that necessitated further testing. However, as a result of this investigation, the failure mode has been better characterised, allowing more informed choices to be made on the design and manufacturing variables that may influence the failure mechanism. In addition, a repeatable test design is now available, alongside a regression model for the baseline condition, both providing the basis for quantitative conclusions to be made.

Similarly, now the main effects on the location and dispersion of NCV seat wear and injector drift are understood, further testing can be conducted with an increased proportion of control samples. Such testing will allow for the hypothesised composition of wear mechanisms to be tested further, providing additional insight into both sliding wear when the valve is closed, and the components of wear associated with valves actuation.

With regards to injector drift, significant dispersion has been observed in the results for all sample populations that cannot be fully attributed to any of the usage variables considered. One possible explanation of this dispersion could be through variations in the samples associated with their manufacture. Interrogation of the manufacturing data associated with the samples used in this investigation could provide additional insight into the relationship between injector drift and variations associated with production, if any. Furthermore, a detailed understanding of the 0-hour condition, with regard to component geometries, of each sample tested in future investigations, either using samples typical of high-volume production, or samples built to an alternative design specification, will likely prove beneficial.

Another possible explanation of the dispersion exhibited is the effects of other variables on the failure mode, either in the form of noise variables, or co-variates. An experimental design suitable for screening, such as the Plackett–Burman designs, could be used to screen a wider number of variables for significant effects on dispersion. An alternative explanation for the dispersion exhibited in the injector performance metrics, is the contribution of other failure modes within the injector. A comprehensive inspection and measurement of the injector components may yield appropriate evidence to support that hypothesis. Similarly, testing the NCV assemblies for each sample with the same, reference nozzle assembly, will de-couple any nozzle related effects on performance.

One goal of such further investigations would perhaps be to be able to attribute the effects of NCV seat wear, and other failure modes, to specific changes in the injector performance. If the results of a performance test could be interrogated in such a way as to attribute likely magnitudes of different failure modes, significant resource savings could be realised in product development programmes through no longer having to dissemble and fully inspect and measure each sample, and instead, be able to monitor failure mode progression over time. Combined with closed loop monitoring and control of future DFI systems, it may be possible to significantly improve robustness to performance change over time.

While the generalised model for injector drift provides an approximation of the effects of different system operating pressures on injector performance over time, it doesn't yet characterise the effects of cyclic duty cycles. The model is derived from tests conducted with steady state rail pressures, but usage duty cycles will vary across the entire operating range, with differing residencies based on application, as shown in §8.4. Further investigation will be required to fully characterise the effects of cyclic pressures, using tests, and corresponding performance inspection intervals, that step to different rail pressures to study the cumulative effects of both increasing, and decreasing rail pressures. Another gap concerning the regression model is the assumption of its asymptotic nature. The model is fitted using data up to 1000hours only, and while a possible asymptotic relationship can be observed, additional data will be required to test that assumption.

With respect to the usage variables identified by the expert panel as being of high significance to NCV seat wear and injector drift, knowledge gaps remain around the influence of fuel properties after the empirical investigation. While fuel type was used as a design variable in the DoE study, no significant effect was observed on the location response of either NCV seat wear or injector drift. As discussed in §7.3.2 and §7.3.4, this is a compound effect of the main hydraulic properties being impractical to isolate, and fuel degradation with recirculation. Further research and iterative empirical investigation will be required to better characterise the effects of fuel properties, if any, on NCV seat wear and injector drift.

The final remaining gap concerns the influence of hard particle debris contamination. Since all treatment combinations in the DoE used the same level of hard particle debris contamination, there is a knowledge gap existing around the influence of both the debris damage itself and any subsequent flow erosion damage to the valve seats. Further investigation, both through interrogation of existing test results, and through further designed experiments, will be required to fully characterise the effect, if any, of this usage variable.

9.6 Discussing the empirical investigation as a part of the overall methodology

Within the overall framework of this research, the empirical investigation has generated new knowledge through designed experiments, and statistically meaningful comparison of sample populations. The DoE programme has proven an effective means for further reducing the system model of the failure mode generated through the expert elicitation study.

The failure mode used in the case study had been the subject of numerous prior empirical studies, and used no new samples, test methods, or test conditions than had previously been used. Those prior studies were resource intensive, lead through the project teams, involving numerous engineers, test rigs, and samples. However, the test designs were largely un-designed, with relatively poor control of variations in test samples and test factors, often changing multiple variables simultaneously, without an understanding of factorial effects. When this approach to test design is combined with the dispersion in results also exhibited in this programme, no sufficiently rigorous conclusions could be made.

By using DoE methodology, specifically a fractional factorial design with a centre point, this empirical test programme has provided the requisite level of statistical confidence to draw conclusions on the relationship between usage variables and NCV seat wear, while representing a resource sensitive approach. However, it is important to be mindful of the iterative nature of empirical investigations: the need for further investigations has been established, but this is to be considered as part of the ongoing process.

9.7 Discussing the methodology employed in this case study with respect to the original industrial motivations

The original motivation of this case study was to identify the usage variable(s) that most directly influence NCV seat wear and injector performance drift such that an accelerated test could be designed. This motivation is reflected in the data analysis methodology employed, where the focus is on identifying, and modelling, only the most significant variable(s). This motivation can be reflected in the Causal Loop Diagrams generated as part of the Expert Elicitation study, which through a process of reduction, guided by the judgement of the expert panel, represented only the most significant variables associated with the failure mode, including usage variables.

Furthermore, considering the original motivation, the results of the Experimental Design employed in the empirical investigation were used to determine which of the design factors were statistically significant, and separable from experimental error or sample-to-sample variance. In this case study, only the most significant variable was then utilised in the regression modelling element while a more classical methodology would have incorporated all the design factors as variables within the regression modelling. These methodological choices are in line with the motivations of the case study, allowing usage variables to be identified for which a significant effect could be observed, providing a means to design accelerated tests, and allowing for a model to be generated that can be used to compare populations using only limited data.

Considering the results of the empirical investigation, the Causal Loop Diagrams then represent an effective, and elegant means for communicating how variance in usage influences the failure mode, as presented in §9.3.2. The CLD presented in Figure 127 demonstrates that while only one usage variable has been demonstrated as significantly influencing the failure mode, it does so in a complex manner, potentially covarying with a number of other parameters, influencing potential robust design solutions, and the design of the accelerated tests that could be used to validate them.

9.8 Sustainable interventions

With regards to characterisation of the failure mode, the final stage of the process, itself outside of the scope of this research, was to use the characterisation of usage variables to develop both robust design solutions to the failure mode, and to develop an accelerated test through which those robust design solutions can be validated. The knowledge derived from this process should was then integrated into the modelling and simulation of future products, and as such, contributes in ensuring the NPD process is moving towards VPE.

This research has established a method with Delphi Technologies for structuring failure mode characterisation that is sensitive to the complex sociotechnical context of the organisation and its products. This method has been incorporated into the DFR tools available within the organisation, used to manage the resolution of product failures, and training in the associated methods has been made available for the appropriate engineering resources of the business.

9.9 Possible generalisation of the method used in the case study

Through a combination of expert elicitation, experimental design, and regression modelling, the application of the FMC method utilised in this thesis has identified a usage variable as having a statistically significant effect on the level of both injector performance change over time, and NCV seat wear. This case study provides significant empirical certainty to a problem that had previously presented uncertainty, while doing so in a resource sensitive manner, through remote application of the Delphi Method, and Fractional Factorial experimental design. In doing so, the knowledge generated allows for robust design solutions to be developed, alongside statistically driven accelerated tests. As firms move towards increased levels of VPE in NPD, tools such as this FMC method best allow distributed tacit knowledge to be combined and codified into numerical models of complex problems.

In doing so, the method used in this case study can be characterised in a number of manners. Firstly, in the context of significant sociotechnical complexity, with multiple prior investigations by numerous actors without the business, the method can be characterised by its ***purposeful planning***. This can be observed in the selection and design of appropriate methodologies to best structure subsequent investigations, while making use of existing knowledge in a resource sensitive manner.

Secondly, this method can be characterised through its ***combination of new and existing knowledge***, with structured interventions observed to rationalise, contextualise, and test the understanding of the failure mode with respect to usage.

Thirdly, this method can be characterised through its emphasis on ***reflection***, with significant time resource spent understanding, analysing, and validating the results of each stage of the FMC method before progressing.

Finally, this method can be characterised through its emphasis on ***iteration***, both within the individual elements of the method, in the multi-step nature of the method itself, and in recognition that the problem statement could be identified as ‘wrong’ at any time.

Furthermore, the same emphases can be seen in the individual methodological elements of this FMC method. The Delphi Method for expert elicitation can be characterised in part by its reflective and iterative nature. Similarly, one the keys to successful implementation of experimental design methods is to focus resource of the planning and analysis of experiments, while recognising that the experimental process will likely be iterative in nature. Regression Modelling is also reflective and iterative in nature and presents a direct means for combining new and existing knowledge into numerical models.

The NPD process itself is iterative, with product designs evolving over time. However, within lean product development, it is meaningful planning and reflection that is often neglected. It is suggested that the FMC method presented in this thesis represents a suitable structure for incorporating reflection and iteration into the problem-solving process.

9.10 Generalisation of the FMC method

While several existing models could be suitable to describe the Failure Mode Characterisation method presented in this thesis in a generalised way, it is suggested that a composite model, while ensuring relative simplicity, may best reflect the emphases of this method. As the method brings together aspects of both knowledge management, and more classical engineering, as two does the generalisation. Considering knowledge management, the Nonaka & Takeuchi SECI model is perhaps one of the predominant, presenting a 4-stage cycle of:

Socialisation – Externalisation – Combination – Internalisation

Considering, the more classical engineering domain, the PDCA Cycle, or Deming wheel, is a predominant model for quality improvement, presenting a 4-stage cycle of:

Plan – Do – Check – Act

The SECI and PDCA models can be aligned, such that the externalisation of knowledge aligns with the planning element of problem solving; combination of knowledge is aligned with generation of new knowledge through doing; internalisation and reflection of knowledge is aligned with checking; and socialisation of knowledge is aligned with acting.

Plan & Externalise – Do & Combine – Check & Internalise – Act and Socialise

Considering the FMC method presented in this thesis, it can be considered as 3 separate 'cycles', with the outcome of the 1st cycle informing the beginning of the 2nd, and so forth. Considering each of those cycles, the method associated with this thesis can be visualised in terms of both the SECI and PDCA models. The Expert Elicitation study discussed in Chapter 6 can be visualised as in Figure 131, where the design, application, and modellisation of the Delphi Study can be expressed as a PDCA/SECI cycle.

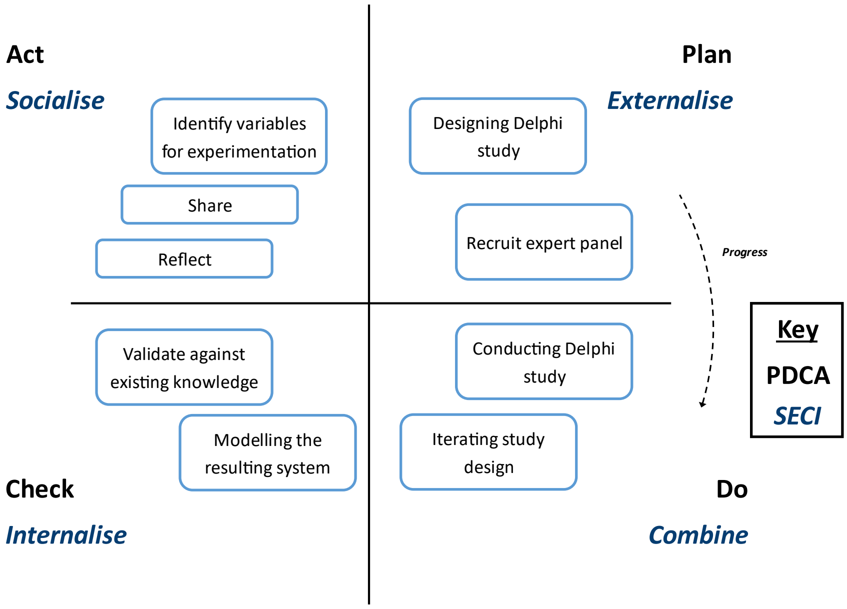


Figure 131: Visualisation of Delphi Study as a PDCA/SECI cycle

The empirical study using experimental design methodology discussed in Chapter 7 can then be visualised as in Figure 132, where outputs of the preview cycle are used in the design, execution, and analysis of the experiment can be expressed as a PDCA/SECI cycle.

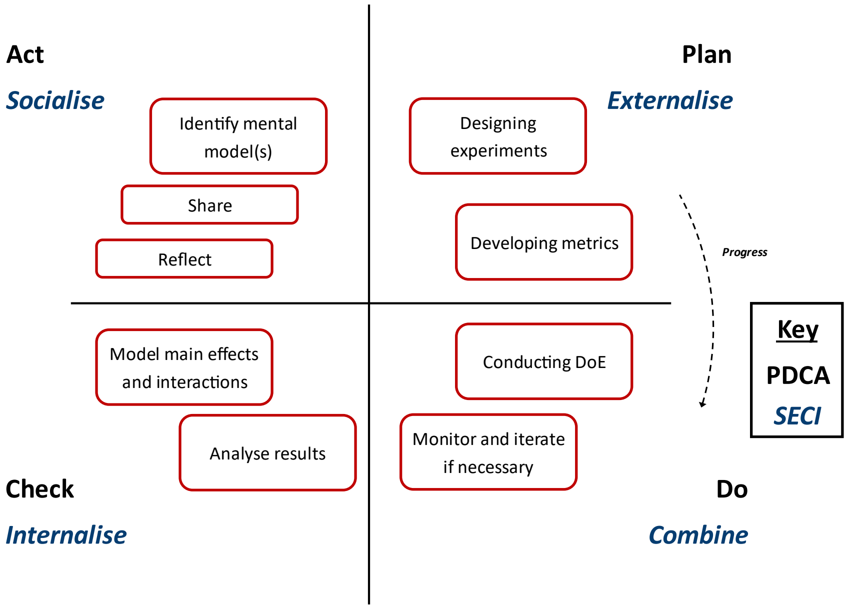


Figure 132:Visualisation of empirical study as a PDCA/SECI cycle

The regression modelling of the failure mode discussed in Chapter 8 can then be visualised as in Figure 133, where outputs of the preview cycle are used in the design, development, and analysis of the numerical models can be expressed as a PDCA/SECI cycle.

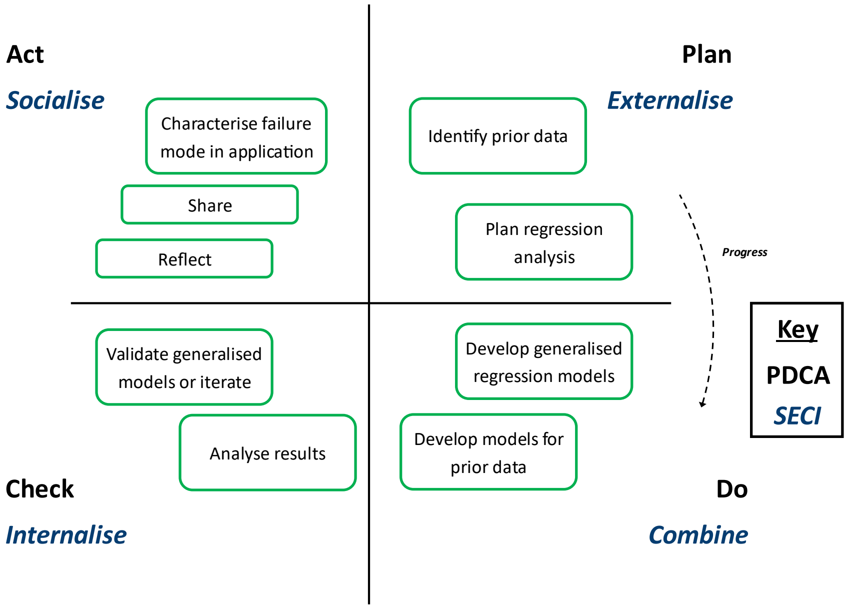


Figure 133: Visualisation of regression modelling as a PDCA/SECI cycle

Considering then the whole method, the three separate cycles presented can be presented using a combination of a Spiral Model, and the PDCA & SECI cycles, where each completed cycle informs the next, as shown in Figure 134.

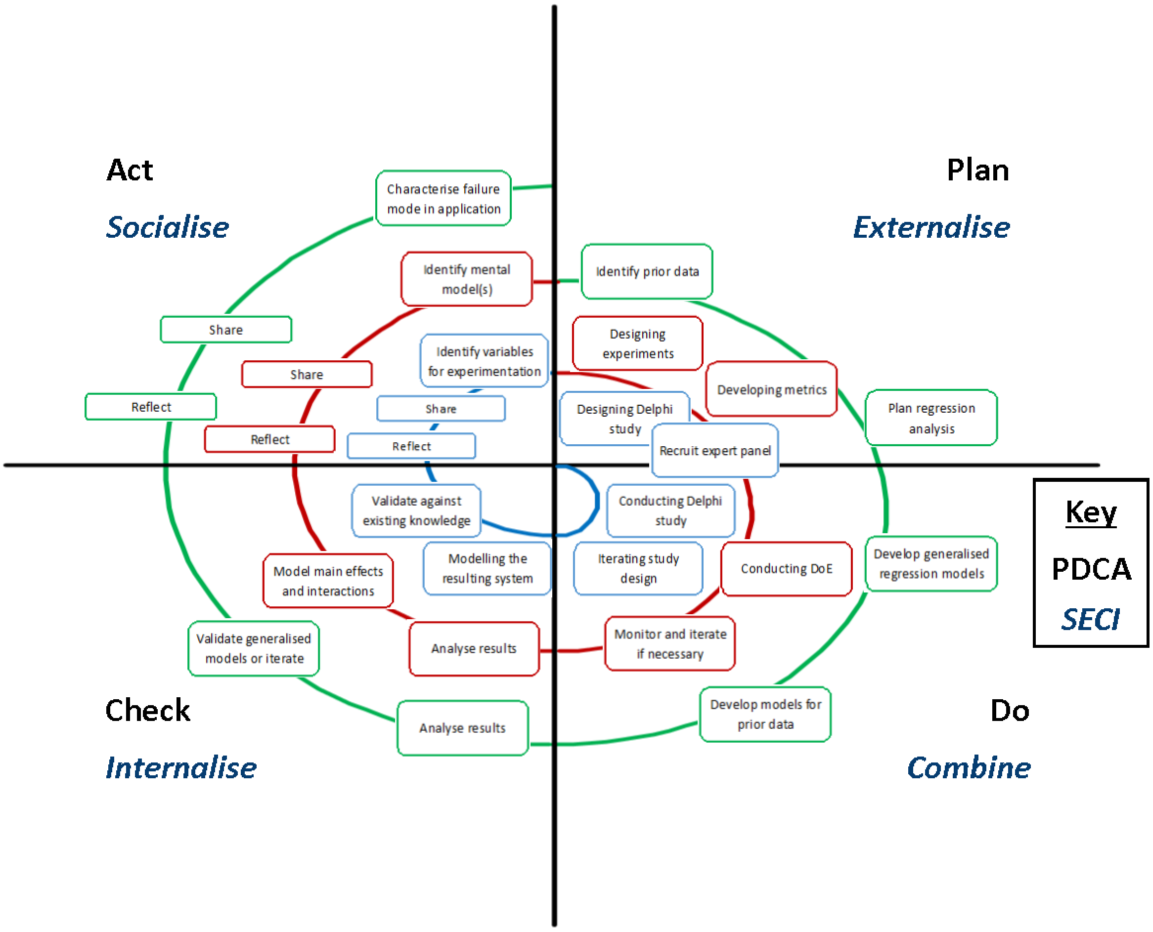


Figure 134: Visualising the 3 PDCA/SECI cycles as a spiral

This model visualises the specific methodological choices that were presented for this case study, but it is acknowledged that the Delphi Method may not always be the most appropriate form of expert elicitation, fractional factorial designs may not always be the most suitable experimental design, and numerical regression modelling may not always be possible or desirable. As such, the method can be generalised further, considering instead the three cycles of:

Expert Judgement Elicitation – Experimentation – Modellisation

It is anticipated that the specific methods used for each cycle would be selected to best balance rigour and other considerations for each application of the FMC method.

Furthermore, the model itself can be generalised. Rather than express the cycle in terms of both the PDCA cycle and the SECI cycle, specific language can be presented that emphasises the key motivations of the FMC method presented in this thesis.

Combining the 'Plan' of PDCA and the 'Externalise' of SECI, it is suggested that 'Design & Plan' represents the sharing of a problem, and purposeful planning and design of the separate elements of this FMC method, emphasising how resource should be focused upfront in each cycle.

Combining the 'Do' of PDCA, and the 'Combine' of SECI, it is suggested that 'Combine and create' represents the focus on combining existing knowledge with newly generated knowledge associated with each element of this FMC method, either through elicitation, experimentation, or modelling.

Combining the 'Check' of PDCA and 'Internalise' of SECI, it is suggested that 'Reflect' emphasizes the validation, analysis, and comparison of new, or newly combined knowledge, against prior knowledge, associated with this FMC method.

Finally, combining the 'Act' of PDCA and the 'Socialise' of SECI, it is suggested that 'consolidate' represents the process through which it is decided to continue or re-iterate the cycle, while new knowledge is shared, and presented, or characterised, in a form suitable for meaningful next steps.

While the 'Consolidate' section of the model explicitly presents a decision point for iteration, it is suggested that at all points in the method, with suitable deference to reflection, the practitioner should consider the suitability of the problem statement and methods selected, and the representativeness of the knowledge combined and generated, with the option to iterate within a cycle, or back to the beginning of the FMC method itself. Rather than indicating deficiencies in methodological choices or in execution, this desire to iterate may be the result of the process creating new knowledge that changes the worldview of the problem, perhaps identifying that the 'wrong' problem has been identified.

The generalised model for FMC is visualised in Figure 135.

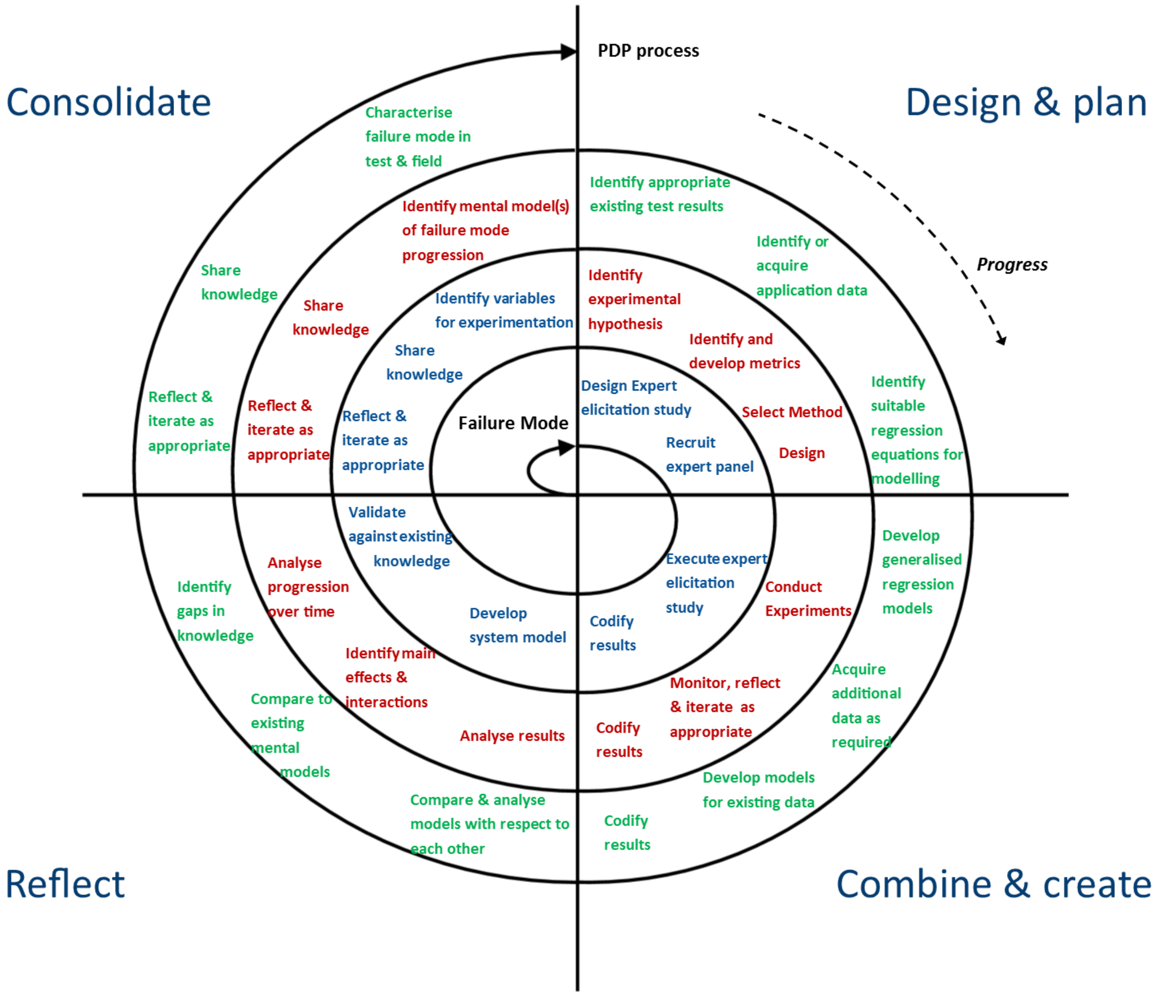


Figure 135: Generalised FMC method

The model presented presents a simple model for practitioners, while emphasising the purposeful planning, combination of new and existing knowledge, reflection, and iteration associated with the FMC method presented in this thesis. However, by not oversimplifying the process into a linear, step-by-step model, this model in turn represents a means for structuring interventions in a complex sociotechnical contest.

9.11 Discussion of possible metrics for FMC methods

The method presented in this research is proposed to be a more resource sensitive, and efficient means of characterise product failure modes than the extant methods used within Delphi Technologies and more generally within the industry. However, to prove or disprove that proposal, comparison of appropriate metrics would have to be performed in a longitudinal study.

Such a study is outside of the scope of this research but would require only existing metrics. To assess the resources required to characterise the failure mode, the proposed metrics are listed in Table 32.

Element of characterisation process	Metric
Delphi Study	Expert engagement (total hours)
	Facilitation (total hours)
System Modelling	Facilitation (total hours)
	Stakeholder engagement (total hours)
Empirical Investigation	Test planning (total hours)
	Samples required (number & cost)
	Test rig requirements (number & total hours)
	Test rig support (total hours)
	Fuel usage (total cost)
	Product characterisation testing (total cost)
	Product inspection & measurement (total hours)
	Test analysis (total hours)
Regression modelling	Regression modelling (total hours)
	Regression analysis (total hours)
Overall	Time to characterise failure mode

Table 32: Proposed metrics for a case study of FMC method

To assess the longitudinal, effectiveness of the method, the proposed metrics are listed in Table 33.

Metric
Subsequent related engineering changes (number of and cost to implement)
Warranty performance (number of, liability of, and cost of)
Customer satisfaction (complaints and repeat business)
End user perception (reliability rankings & industry press reports)

Table 33: Proposed longitudinal metrics for the effectiveness of the FMC method

An improved method would therefore be one that requires less resources to initially conduct, while also reducing the overall engineering costs associated with launching the product, its warranty performance in application, and brand reputation.

9.12 Summary

This chapter has discussed the case study of Failure Mode Characterisation presented in this thesis, in terms of its design, execution, and results. The findings of the case study are discussed with regards to the sociotechnical context of the problem, and the knowledge gaps that remain were presented. Finally, the method used in the case study was discussed, and a generalised form was proposed as a simple tool for practitioners, with metrics suggested for how its effectiveness could be evaluated.

The next Chapter will conclude this thesis with regards to the original Research Hypothesis and questions.

Chapter 10 Conclusions

10.1 Introduction

This Chapter will conclude the thesis. The Research Hypothesis and Questions presented in Chapter 1 will be revisited in context of the outcomes of the case study, with the relevant conclusions identified. Opportunities for future search will also be presented.

10.2 A systematic method for failure mode characterisation in a complex sociotechnical context

This thesis has presented an industrial case study of FMC in a complex Sociotechnical context. The Research Hypothesis presented in Chapter 1 was:

> *In a complex sociotechnical organisation engaged in the development of Fuel Injection Systems, can a systematic method for failure mode characterisation be developed that creates new knowledge of a failure mode subject to previous, unsuccessful investigations, and can that method be generalised for organisations engaged in New Product Development.*

The following sections will identify the conclusions associated with each Research Question that were proposed to test that hypothesis.

10.2.1 Research Question 1: How can expert judgement relevant to DFI control valve seat wear be elicited in an effective, and resource sensitive manner?

In Chapter 6, this thesis presented a case study of the application of the Delphi Method as an Expert Elicitation method, with a geographically distributed expert panel (§6.3). Drawing from knowledge of previous investigations, the Delphi Method resulted in a group-led definition of the failure mode (§6.4.2), and the identification of 35 variables, and their interactions, associated with the design, manufacture, and usage of the products that influenced the severity of the failure mode (§6.4.3). The Delphi Study also resulted in the attribution of significances to each variable by the panel, with 11 variables reaching consensus as being of little to no significance, while informed disagreement was met on the significance of 14 variables (§6.4.3). The study also elicited the expert's panels worldviews of the importance of the failure mode (§6.4.4), and their views on why previous investigations had failed to characterise it (§6.4.5).

The Delphi Method is presented as a resource sensitive, and effective means for eliciting the judgement of the expert panel in this case study (§9.2.1), with positive and constructive feedback received from the panel.

10.2.2 Research Question 2: Can a suitable modelling technique be identified to codify the results of the expert elicitation as a useful boundary object in the NPD process?

In Chapter 6, this thesis presented case study of the use of Causal Loop Diagrams as a means for codifying the expert judgement elicited in the Delphi Study (§6.5). A model of the failure mode was generated firstly using only the qualitative discussion of the panel (§6.5.1), before then being augmented to represent every variable identified in the study (§6.5.2). The model was developed further to include visualisation of the domain (design, manufacture, or usage) associated with each variable, alongside the significances identified by the panel (§6.5.2). Through stakeholder engagement, the model resolution was iterated such as to serve as a useful boundary object while retaining the most significant knowledge (§6.5.2). Finally, Causal Loop Diagrams were also used to capture the views of the panel on why previous investigations had failed, providing support to the intended industrial objectives of this research (§6.5.3).

Causal Loop Diagrams are presented as an effective tool for codifying the results of the expert elicitation process, visualising the failure mode, and the judgement of the expert panel (§9.2.2).

10.2.3 Research Question 3: Does structured empirical investigation represent an effective and resource sensitive means for identifying the significant effects and interactions associated with the usage variables identified through expert elicitation?

In Chapter 7, this thesis presented a case study of the use of Experimental Design methodology in the characterisation of usage variables identified in the Delphi Study as being of potential significant to the failure mode. A review of the variables where either consensus was met as being of high significance, and the variables where informed disagreement was reached, identified 3 usage variables suitable for empirical investigation (§7.3). After some iterative experimentation and associated learning, a half fractional factorial design was selected (§7.3.6), with 3 factors each at two levels (§7.3.7). For each treatment combination, a total of 12 samples were used, one of which was electronically disconnected to represent a form of a control sample (§7.4.6). The design factors used were: system operating pressure, number of injections, and fuel type. Response metrics were selected, and developed as appropriate, representing both injector performance (§7.5) and physical measurements of wear (§7.6).

A series of experimental trials were conducted, each with injector performance characterisation tests conducted at 0, 500, and 1000 hours, with physical measurements of the valve seats being completed at 1000hours (§7.9). The results for each treatment combination for each response metric were presented (§7.9.1 – §7.9.5), while an analysis of the Experimental Design identified significant main effects and interactions (§7.9.6 & §7.9.7). The System operating pressure variable was found to result in a significant response in both the TI drift performance metric, and the physical measurement of NCV seat wear, while the fuel type variable resulted in a significant response in the MDP timing metric (§7.9.5).

Using the sparsity of effects principle, the number of injections variable was removed through design projection, allowing additional confidence in the conclusion that rail pressure was the only significant usage variable for injector drift (§7.10). Using the same projected design, a single pseudo centre-point was added to allow inference of the linearity of the response associated with rail pressure. This additional test resulted in the conclusion that the rail pressure effect was non-linear, to which a suitable regression model was fitted (§7.11).

The response of the control samples was compared to the response of the sample populations, allowing additional hypothesis to be suggested. The relative response of the control samples, in terms of both injector performance and physical wear, suggested that there was is a component of the failure mode that varies with rail pressure, but is not related to actuation of the NCV. Furthermore, the component of the failure mode that is related to actuation does not appear to vary with rail pressure (§9.3.2).

The Trimmed Integral of Drift metric developed in this thesis is presented as the most suitable performance-based metric of the failure mode, with correlation in directionality and magnitude of response with the physical seat wear, representing a suitable inference for degradation testing (§9.3.2).

The knowledge gaps that remained after the DoE were discussed (§9.5), and the experimental design element of the FMC method was presented as an effective means for providing robust new knowledge of the usage variables that influenced the failure mode (§9.6). The specific methodological choices associated with the analysis of data in this case study were discussed (§9.7).

10.2.4 Research Question 4: Can a generalised model be developed from the empirical results that adequately describes the failure mode with respect to existing knowledge, while serving as the basis for further investigations?

In Chapter 8, this thesis presented a case study of regression modelling of injector drift, expressed using the TI drift metric, using the new knowledge generated by the DoE programme. Using the available injector performance characterisation results for the original 4 treatment combinations, basic regression models were fitted, but were demonstrated to not be suitable based on the existing knowledge of the failure mode (§8.2.1). When the additional low-hour performance characterisation results associated with the pseudo centre-point were considered, alternative regression models could then be considered (§8.2.2). The Michaelis-Menten model was presented as the best approximation of the failure mode, both in terms of statistical correlation, and with reference to the mental model of the failure mode (§8.2.2). Using all the data from the DoE programme, a generalised model for injector drift was presented that describes injector drift over time in terms of the rail pressure of the FIS (§8.2.3). This model was based on several assumptions (§9.4.1), with limitations identified that could be addressed through additional experimental investigation (9.4.2).

The generalised model of drift is then used to validate the results of this research against of DFI of the same design tested on hydraulic test rigs (§8.3). A regression model is developed for the previous test results (§8.3), which is then compared to the generalised model, demonstrating a significantly reduced progression of drift with respect to time. The test cycle associated with the previous results is then presented, demonstrating a 33%-time residency at 2400bar, while all other cycle points were below the range of pressures used in this research. A time-time based transformation was then applied to the previous test results to account for that residency, which was shown to provide a good approximation to the generalised model for an equivalent rail pressure level.

The generalised model of drift was then used to compare three different drive types representative of the end-user application (§8.4.1). The drive types were compared in terms of vehicle speed and rail pressure distributions, demonstration how the different drive types significantly influence the rail pressure of the FIS, and ultimately, the progression of injector drift, in terms of both hours, and cumulative vehicle drive distance.

The drift regression models for the 3 drive types are then compared to the model for the previous cyclic testing, and the steady state testing associated with this research (§8.4.2). Steady state testing is shown to demonstrate the most accelerated drift in terms of hours, while the most severe applications can be seen to demonstrate an acceleration compared to the previous cyclic rig testing. However, when legislation that limits driving time is considered, cyclic rig testing is demonstrated to demonstrate a slight acceleration over the most severe use cases. The potential implications of this knowledge on application warranty was then discussed (§9.4.2). This knowledge is then discussed in terms of the perception that application testing in engine or vehicles is more accelerated than hydraulic test rigs for this failure mode, presenting several possible sociotechnical explanations (§9.4.3).

Finally, the limitations and assumptions associated with the regression modelling were discussed (§9.4), while it was proposed that it represents an effective element of the FMC method, transforming the new knowledge into the basis for the design of accelerated tests, and improved modelling of the failure mode in CAE environments (§9.4).

10.2.5 Research Question 5: Can this method for FMC be generalised as a resource sensitive method for complex sociotechnical contexts?

In Chapter 9, this thesis presented a discussion of the case study, discussing the method used in the case study (§9.9), suggesting it could be characterised: purposeful planning; a combination of existing and new knowledge; reflection; and iteration. It is discussed how such a method for knowledge transforming and generation could be a suitable tool for firms moving towards increased levels of Virtual Product Engineering.

The 3 distinct stages of the method used in the case study are then each shown to be possible to follow a cycle that represents a composite of the SECI knowledge cycle, and the PDCA cycle for product improvement (§9.10). A generalised form of the FMC method is presented, recognising that specific methodological choices would be based on the problem context. The generalised model transforms the 3 stages of the method into a spiral that represents a cycle of: Plan & Design – Combine & Create – Reflect – Iterate. It is suggested that the model represents a simple method for practitioners.

Metrics for a longitudinal study of the effectiveness of the FMC method are then discussed, with regards to both the product engineering resources required, and for its impact expressed in terms of customer requirements (§9.11).

10.3 Further work

With respect to the specific model for the failure mode presented in this thesis, additional empirical investigations can further test the assumptions associated with the model, while addressing some of its identified limitations. By testing outside of the experimental boundaries associated with this research, the model can be expanded whilst also confirming whether additional main effects or interactions can be observed. Furthermore, additional measurement observations at different test durations, both inside and outside of the boundaries associated with this research, can further improve the confidence of the model's predictions. Further investigation is required to understand and fully characterise the effects of cyclic stress profiles, through incorporation of cumulative damage models or similar. Similarly, efforts should also be made to test and improve the model's predictions for vehicle-based progression of the failure mode, acknowledging the challenges associated with doing so. Finally, additional experimental work and/or analysis should focus on the effects on dispersion in results and should be included in the model as appropriate. Effort should then be made to incorporate the knowledge codified in this model into the design of ALT programmes and in modelling the failure mode analytically. Once validated with further data, the model presented, specifically the value of Theta 2 of 200 hours, would allow the use of Accelerated Degradation tests of significantly reduced duration to validate robust design solutions to this failure mode.

With respect to the more generalised characterisation of the NCV seat wear, additional empirical and analytical investigations can address the gaps that exist after this research. Further investigations should expand the characterisation into the variables associated with the design and manufacture of the fuel injector, with complete characterisation of the samples zero-hour geometries, and assembly tolerances in order to understand any additional effects or interactions associated with either the location or dispersion of the failure mode over time. In addition, further component level investigation of the test samples from this research, or additional empirical testing, could further characterise the failure mode, and potentially further inform the relationship between NCV seat wear, injector performance drift, and any other degradation modes in the injector. Additional empirical testing, or analysis of existing results, will further the understanding of the influence, if any, or hard particle contamination on NCV seat wear and/or injector performance drift. Furthermore, additional investigations are required to understand, and characterise the effects of fuel lubrication on the coated seats of the NCV, and how changes in fuel properties, or the condition of the seats, influence NCV seat wear.

Finally, with respect to the generalised FMC method presented as an outcome of this research, further longitudinal studies of applications in complex sociotechnical contexts will test the suitability and efficacy of this method in industry using the metrics proposed in this thesis.

10.4 Summary

Finally, this research sought to characterise the usage variables that influence control valve seat wear in diesel fuel injectors such that representative accelerated tests could be designed, and robust design solutions could be validated. With significant prior investigations through multiple, geographically dispersed engineering teams, this characterisation represented a complex sociotechnical problem. In this case study, the Delphi Method for Expert Judgement Elicitation resulted in a group led definition of the problem, along with identification of the design, manufacturing, and usage variables that influence the failure mode, along with a description of their interactions. Causal Loop Diagrams were then used to generate a model of the failure mode, that represented the complexity of the problem while serving as a useful boundary object. Extensive empirical testing, using Experimental Design Methodology was then used to identify that Rail Pressure was the only usage variable demonstrated to have a statistically significant effect on the response of the physical wear of the control valve, and as inferred through injector performance change over time. Finally, regression modelling was used to generate a generalised model of the failure mode that can be used to predict the response over time for given usage conditions. Further work should focus on further characterising the failure mode, including the design and manufacturing variables identified by the expert panel.

This case study therefore successfully characterised the usage variables that influence control valve seat wear, while the method employed is suggested as a simple model for practitioners.

References

Antony, J. (1998) ‘Some key things industrial engineers should know about experimental design’, Logistics Information Management, 11(6), pp. 386–392.

Antony, J. and Antony, F. J. (2001) ‘Teaching the Taguchi method to industrial engineers’, Work Study, 50(4), pp. 141–149.

Archard, J., (1953) ‘Contact and rubbing of flat surfaces’, Journal of applied physics, 24(8), pp.981-988.

Automotive Council UK (2016) ‘Technology Roadmap 2016: Virtual Product Engineering’, accessed 29th November 2017 <https://www.automotivecouncil.co.uk/wp-content/uploads/sites/13/2017/09/Virtual-Product-Engineering.jpg>.

Automotive Council UK (2017) ‘Technology Roadmap 2017: Commercial and Off-highway Vehicle’, accessed 29th November 2017, <https://www.automotivecouncil.co.uk/wp-content/uploads/sites/13/2017/09/Virtual-Product-Engineering.jpg>.

Bedford, T. J., Quigley, J. L. and Walls, L. A. (2006) ‘Expert elicitation for reliable system design’, Statistical Science, 21(4), pp. 428–450. doi: 10.1214/088342306000000510.

Bernad, S. I. and Susan-Resiga, R. (2012) ‘Numerical Model for Cavitational Flow in Hydraulic Poppet Valves’, Modelling and Simulation in Engineering, 2012. doi: 10.1155/2012/742162.

Bertsche, B. (2008) ‘Reliability in automotive and mechanical engineering: determination of component and system reliability’, Springer Science & Business Media.

Blau, P. J. and Hanft, T. A. (1994) ‘Quantitative measurement of repetitive impact damage on ground silicon nitride surfaces’, Tribology International, 27(2), pp. 109–118.

Blockley, D. I., & Godfrey, P. (2000) ‘Doing it differently: systems for rethinking construction’, Thomas Telford.

Bolger, F. and Wright, G. (2011) ‘Improving the Delphi process: Lessons from social psychological research’, Technological Forecasting and Social Change. Elsevier Inc., 78(9), pp. 1500–1513. doi: 10.1016/j.techfore.2011.07.007.

Booker, J.D., Raines, M., & Swift, K.G. (2001) ‘Designing Capable and Reliable Products’, First Edition, Butterworth-Heineman.

Box, G. E., & Behnken, D. W. (1960) 'Some new three level designs for the study of quantitative variables', Technometrics, 2(4), 455-475.

Braglia, M., Fantoni, G., & Frosolini, M. (2007) 'The house of reliability', International Journal of Quality & Reliability Management, 24(4), 420-440.

Burgman, M. a. et al. (2011) 'Expert status and performance', PLoS ONE, 6(7), pp. 1–7. doi: 10.1371/journal.pone.0022998.

Campean, F., Wright, S., Brunson, D., and Mistry, S. (2002) 'Vehicle Foresight : Customer Correlation of Engine Components Tests Using Life Prediction Modelling', SAE World Congress - Foresight Vehicle Technology: Consumer Driven Design, Manufacturing, Supply Chain, and Purchasing, (724).

Campean, I. F., Henshall, E. and Rutter, B. (2013) 'Systems Engineering Excellence Through Design : An Integrated Approach Based on Failure Mode Avoidance', SAE Int. J. Mater. Manf., 6(3), pp. 389–401. doi: 10.4271/2013-01-0595.

Carlile, P. R. (2002) 'A Pragmatic View of Knowledge and Boundaries: Boundary Objects in New Product Development', Organization Science, 13(4), pp. 442–455.

Cavana, R. Y. and Mares, E. D. (2004) 'Integrating critical thinking and systems thinking: From premises to causal loops', System Dynamics Review, 20(3), pp. 223–235. doi: 10.1002/sdr.294.

Chang, H., Tsai, M. and Tsai, C. (2011) 'Complex organizational knowledge structures for new product development teams', Knowledge-Based Systems. Elsevier B.V., 24(5), pp. 652–661. doi: 10.1016/j.knosys.2011.02.003.

Chang, K., Lee, S. and Won, I. (2005) 'KMPI: measuring knowledge management performance', Information & Management, 42, pp. 469–482. doi: 10.1016/j.im.2004.02.003.

Collis, J. and Hussey, R., (2013) 'Business research: A practical guide for undergraduate and postgraduate students', Macmillan International Higher Education.

Condra, L.W. (1993) 'Reliability Improvement with Design of Experiments', Second Edition, New York: Marcel Dekker, Inc.

Cooke, R. M. (2013) 'Validating Expert Judgement with the Classical Model', Experts and Consensus in Social Science, (2008), pp. 1–18.

Cooper, L. P. (2003) ‘A research agenda to reduce risk in new product development through knowledge management : a practitioner perspective’, Journal of Engineering and Technology Management, 20, pp. 117–140. doi: 10.1016/S0923-4748(03)00007-9.

Creswell, J.W., (2003) ‘Research Design: Qualitative, Quantitative, and Mixed Methods Approaches’, Second Edition, Sage Publications, London.

Day, J. and Bobeva, M. (2005) ‘A Generic Toolkit for the Successful Management of Delphi Studies’, Electronic Journal of Business Research Methods, 3(2), pp. 103–116.

Delbecq, A.L., Van de Ven, A.H. and Gustafson, D.H. (1975) ‘Group techniques for program planning: A guide to nominal group and Delphi processes’ (pp. 83-107).

Doikin, A. et al. (2018) ‘Impact of Duty Cycle on Wear Progression in Variable-displacement Vane Oil Pumps’, Procedia Manufacturing. Elsevier B.V., 16, pp. 115–122. doi: 10.1016/j.promfg.2018.10.170.

Donohoe, H. M. and Needham, R. D. (2009) ‘Moving Best Practice Forward : Delphi Characteristics, Advantages, Potential Problems and Solutions’, International Journal of Tourism Research, 437 (December 2008), pp. 415–437.

Elsayed, E. A. (2012) ‘Overview of reliability testing’, IEEE Transactions on Reliability, 61(2), pp. 282–291. doi: 10.1109/TR.2012.2194190.

Erto, P. and Giorgio, M. (2002) ‘Assessing high reliability via Bayesian approach and accelerated tests’, Reliability Engineering and System Safety, 76, pp. 301–310.

Escobar, L. a. and Meeker, W. Q. (2006) ‘A Review of Accelerated Test Models’, Statistical Science, 21(4), pp. 552–577. doi: 10.1214/088342306000000321.

Firka, D. (2011) ‘Statistical , technical and sociological dimensions of design of experiments’, The TQM Journal, 23(4), pp. 435–445. doi: 10.1108/17542731111139509.

Ford, D. and Sterman, J. D. (1998) ‘Expert Knowledge Elicitation to improve formal and mental models’, System Dynamics Review, 14(4), pp. 309–340.

Fricke, R.W. and Allen, C., (1993) ‘Repetitive impact wear of steels’, Wear, 162, pp.837-847.

Geist, M. R. (2010) ‘Using the Delphi method to engage stakeholders: A comparison of two studies’, Evaluation and Program Planning, 33(2), pp. 147–154. doi: 10.1016/j.evalprogplan.2009.06.006.

Gill, F. J. et al. (2013) 'Using a web-based survey tool to undertake a Delphi study : Application for nurse education research', YNEDT. Elsevier Ltd, 33(11), pp. 1322–1328. doi: 10.1016/j.nedt.2013.02.016.

Goossens, L. H. J. et al. (2008) 'Fifteen years of expert judgement at TUDelft', Safety Science, 46, pp. 234–244. doi: 10.1016/j.ssci.2007.03.002.

Gordon, T. and Pease, A. (2006) 'RT Delphi: An efficient, "round-less" almost real time Delphi method', Technological Forecasting and Social Change, 73(4), pp. 321–333. doi: 10.1016/j.techfore.2005.09.005.

Graham, M.S., Crossley, S., Harcombe, T., Keeler, N. and Williams, T., (2014) 'Beyond Euro VI-Development of A Next Generation Fuel Injector for Commercial Vehicles' (No. 2014-01-1435), SAE Technical Paper.

Haraldsson, H. (2000) 'Introduction to system and causal loop diagrams', System Dynamic Course, Lumes, Lund University, …, (January), pp. 1–33. Available at: http://dev.crs.org.pl:4444/rid=1244140954250_1167059429_1461/Introduction to Systems and Causal Loop Diagrams.pdf.

Hardy, C.S., Hudson, C.M., Harper, M.S. and Ainsworth, D.M. (2012) 'Pressure is nothing without control: Evolution of control valve design', In Fuel systems for IC engines (pp. 115-128).

Henshall, E., Campean, I. F. and Rutter, B. (2014) 'A Systems Approach to the Development and Use of FMEA in Complex Automotive Applications', SAE International Journal of Materials and Manufacturing, 7(2), pp. 280–290. doi: 10.4271/2014-01-0740.

Hockman, K. K., & Jenkins, M. W. (1994) 'Design of experiments: Key to competitive R&D', Quality, 33(6), 47.

Ilzarbe, L. et al. (2008) 'Practical Applications of Design of Experiments in the Field of Engineering: A Bibliographical Review', Quality and Reliability Engineering International, 24(February), pp. 417–428. doi: 10.1002/qre.

Kuhn, T., (1962) 'The structure of scientific revolutions', Chicago: Univ. Press, Chicago.

Lacey, P., Gail, S., Kientz, J.M., Milovanovic, N. and Gris, C., (2012) 'Internal fuel injector deposits'. SAE International Journal of Fuels and Lubricants, 5(1), pp.132-145.

Lacey, P.I., Gunsel, S., De La Cruz, J. and Whalen, M.V., (2001) 'Effects of high temperature and pressure on fuel lubricated wear' (No. 2001-01-3523), SAE Technical Paper.

Landeta, J. (2006) 'Current validity of the Delphi method in social sciences', Technological Forecasting and Social Change, 73, pp. 467–482. doi: 10.1016/j.techfore.2005.09.002.

Lane, D. C. (2008) 'The emergence and use of diagramming in system dynamics: A critical account', Systems Research and Behavioral Science, 25(1), pp. 3–23. doi: 10.1002/sres.826.

Lewis, R. (2007) 'Short communication A modelling technique for predicting compound impact wear', Wear, 262, pp. 1516–1521. doi: 10.1016/j.wear.2007.01.032.

Lewis, R. and Dwyer-Joyce, R. (2002) 'Wear of diesel engine inlet valves and seat inserts', Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering, 216, pp. 205–216. doi: 10.1243/0954407021529048.

Lewis, R., Dwyer-Joyce, R.S. and Josey, G., (1999) 'Investigation of wear mechanisms occurring in passenger car diesel engine inlet valves and seat inserts' (No. 1999-01-1216), SAE Technical Paper.

Lillington, R., Chetwynd, D. G. and Jones, R. P. (2015) 'Characterising heavy-duty fuel systems : a critique of industry practice and challenges', Proceedings of the Institution of Mechanical Engineers , Part D : Journal of Automobile Engineering, 229(3), pp. 391–403. doi: 10.1177/0954407014540899.

Liu, P., Chen, W. and Tsai, C. (2005) 'An empirical study on the correlation between the knowledge management method and new product development strategy on product performance in Taiwan' s industries', Technovation, 25, pp. 637–644. doi: 10.1016/j.technovation.2003.11.001.

Lloyd, P. (2015) Drift Task Force. Internal Delphi Technologies report. Unpublished.

Logothetis, N. (1994) 'Managing for total quality: from Deming to Taguchi and SPC', Prentice Hall.

Mackenzie, N. and Knipe, S. (2006) 'Research dilemmas: Paradigms, methods and methodology', Issues in Educational Reseae, 16, pp. 1–11.

Meeker, W. Q. and Escobar, L. A. (1993) 'A Review of Recent Research and Current Issues in Accelerated Testing', International Statistical Review, 61(1), pp. 147–168.

Meeker, W. Q., Sarakakis, G., & Gerokostopoulos, A. (2013) 'More pitfalls of accelerated tests', Journal of Quality Technology, 45(3), 213-222.

Mingers, J. (2011) 'Soft OR comes of age-but not everywhere!', Omega. Elsevier, 39(6), pp. 729–741. doi: 10.1016/j.omega.2011.01.005.

Mingers, J. and Brocklesby, J. (1997) 'Multimethodology: Towards a framework for mixing methodologies', Omega, 25(5), pp. 489–509. doi: 10.1016/S0305-0483(97)00018-2.

Mingers, J. and Rosenhead, J. (2004) 'Problem structuring methods in action', European Journal of Operational Research, 152(3), pp. 530–554. doi: 10.1016/S0377-2217(03)00056-0.

Montgomery, D.C., (2009) 'Design and analysis of experiments', Seventh Edition, John Wiley & Sons, Inc..

Morecroft, J. D. W. (1982) 'A critical review of diagramming tools for conceptualizing feedback system models', Dynamica, pp. 20–29.

Nelson, W. (1990) 'Accelerated Testing: Statistical Models, Test Plans, and Data Analyses', John Wiley & Sons.

Nonaka, I. and Takeuchi, H. (1995) 'The Knowledge-Creating Company', Oxford University Press.

O'Connor, P. D. (2001) 'Test engineering: a concise guide to cost-effective design, development and manufacture', Chichester: Wiley.

Okoli, C. and Pawlowski, S. D. (2004) 'The Delphi method as a research tool: an example, design considerations and applications', Information & Management, 42(1), pp. 15–29. doi: 10.1016/j.im.2003.11.002.

Patton, M. Q. (1997) 'Utilization focused evaluation: The new century text'. Thousand Oaks, CA: Sage.

Pignatiello, J. J. and Ramberg, J. S. (1991) 'Top Ten Triumphs And Tragedies Of Genichi Taguchi', Quality Engineering, 4(2), pp. 37–41. doi: 10.1080/08982119108918907.

Prvan, T. and Street, D. J. (2002) 'An annotated bibliography of application papers using certain classes of fractional factorial and related designs', Journal of Statistical Planning and Inference, 106, pp. 245–269.

Ragab, M. A. F. and Arisha, A. (2013) 'Knowledge management and measurement: a critical review', Journal of Knowledge Management, 17(6), pp. 873–901. doi: 10.1108/JKM-12-2012-0381.

Richardson, G. P. (1986) 'Problems with causal-loop diagrams', System Dynamics Review, 2(2), pp. 158–170.

Rosenhead, J. (1996) 'What's the problem? An introduction to problem structuring methods.', Interfaces, 26(6), pp. 117-131.

Rothwell, W. J. and Sullivan, R. (2005) 'Practicing Organization Development', 2nd Ed, Pfeiffer.

Rowlands, H., Antony, J. and Knowles, G. (2000) ‘An application of experimental design for process optimisation’, The TQM Magazine, 12(2), pp. 78–83.

Sainter, P. et al. (2000) ‘Product Knowledge Management Within Knowledge-Based Engineering Systems’, in Proceedings of DETC’00.

Sarakakis, G., Gerokostopoulos, A., & Mettas, A. (2011), ‘Special topics for consideration in a design for reliability process’, In Reliability and Maintainability Symposium (RAMS), 2011 Proceedings-Annual (pp. 1-6). IEEE.

Sarakakis, G., Lassar, N. and Chowdhury, S. (2013) ‘Designing a System to Manage Field Reliability’, in Annual Reliability and Maintainability Symposium.

Shabi, J. and Reich, Y. (2012) ‘Advanced Engineering Informatics Developing an analytical model for planning systems verification , validation and testing processes’, Advanced Engineering Informatics. Elsevier Ltd, 26(2), pp. 429–438. doi: 10.1016/j.aei.2012.02.001.

Sherman, J. D., Berkowitz, D. and Souder, W. E. (2005) ‘New Product Development Performance and the Interaction of Cross-Functional Integration and Knowledge Management’, The Journal Of Product Innovation Management, 22, pp. 399–411.

Skulmoski, G. J. and Hartman, F. T. (2007) ‘The Delphi Method for Graduate Research’, Journal of Information Technology Education, 6(1), pp. 1–21. doi: 10.1.1.151.8144.

Sokovic, M., Pavletic, D. and Pipan, K. K. (2010) ‘Quality Improvement Methodologies – PDCA Cycle, RADAR Matrix, DMAIC and DFSS Industrial management and organisation Industrial management and organisation’, Journal of Achievements in Materials and Manufacturing Engineering, 43(1), pp. 476–483.

Steiner, S. H., MacKay, R. J., & Ramberg, J. S. (2007) ‘An overview of the Shainin System™ for quality improvement’, Quality Engineering, 20(1), 6-19.

Suh, N.P. and Sridharan, P., (1975) ‘Relationship between the coefficient of friction and the wear rate of metals’, Wear, 34(3), pp.291-299.

Tanco, M. et al. (2008) ‘How is Experimentation Carried Out by Companies? A Survey of Three European Regions’, Quality And Reliability Engineering International, 24(August), pp. 973–981. doi: 10.1002/qre.

Thomke, S. and Bell, D. E. (2001) ‘Sequential Testing in Product Development’, Management Science, 47(2), pp. 308–323. doi: 10.1287/mnsc.47.2.308.9838.

Thomke, S.H., (2003) 'Experimentation matters: unlocking the potential of new technologies for innovation'. Harvard Business Press.

Vaccaro, A., Parente, R. and Veloso, F. M. (2010) 'Knowledge Management Tools , Inter-Organizational Relationships , Innovation and Firm Performance', Technological Forecasting & Social Change. Elsevier Inc., 77(7), pp. 1076–1089. doi: 10.1016/j.techfore.2010.02.006.

Vassiliou, P. and Mettas, A. (2003) 'Understanding Accelerated Life-Testing Analysis', in Annual Reliability and Maintainability Symposium.

Venkatasubramanian, V. (2009) 'DROWNING IN DATA : Informatics and Modeling Challenges in a Data-Rich Networked World', AIChE Journal, 55(1). doi: 10.1002/aic.

Wu, H. and Wu, Z. (2012) 'Combustion characteristics and optimal factors determination with Taguchi method for diesel engines port-injecting hydrogen', Energy. Elsevier Ltd, 47(1), pp. 411–420. doi: 10.1016/j.energy.2012.09.027.

Yadav, O. P., Choudhary, N. and Bilen, C. (2008) 'Complex System Reliability Estimation Methodology in the Absence of Failure Data', Quality And Reliability Engineering International, 24(May), pp. 745–764. doi: 10.1002/qre.

Zielenski, M., Downes, P., & Jelbert, D. (2014) 'Adapting Validation Processes to Meet Worldwide Market Demands' (No. 2014-01-1437), SAE Technical Paper.

Appendix I Legislative and customer drivers for Diesel FIS

This Appendix details the market drivers for Diesel FIS systems, drawing from the author's experience, and information in the public domain.

Compliance with Global Emissions Legislations

Commercial Vehicle OEMs, of both on highway and off-road vehicles, are required to meet stringent emission legislations in most markets. Emissions can be either expressed as 'engine out', relating to emissions directly resulting from combustion, or as 'tailpipe', including any Exhaust Aftertreatment Systems (EATS), including Diesel Oxidation Catalysts (DOC), Selective Catalytic Reduction systems (SCR) or Diesel Particulate Filters (DPFs). The legislative emissions tests include any EATS employed in application and are thus considered 'tailpipe'.

The European market, alongside the US market, represents the most challenging emissions requirements, often setting standards that other territories follow at later dates. As such, it is the European legislation that will generally be presented hereby. In addition to more stringent limits for PM and NO_x, Euro VI specifies limits on both Carbon Monoxide (CO) and unburnt hydrocarbons (HC), while including a limit on Particle Number (PN), limiting the total number of particles in the exhaust in order to control the quantity of the smallest particulate emissions associated with respiratory conditions.

The Euro VI legislation also included a change in the engine test procedures used for emissions certification, introducing the two World Harmonised Cycles (WHC). This meant the steady state test changing from the European Steady-State Cycle (ESC) to the World Harmonised Steady-State Cycle (WHSC), and the transient test changing from the European Transient Cycle (ETC) to the World Harmonised Transient Cycle (WHTC). The WHC were derived to be more representative of in-service conditions for the on-highway heavy-duty market, and the WHTC places more emphasis on low load & speed conditions than the previous ETC, with the lower load impacting exhaust gas temperatures and the effectiveness of EATS.

In order to better ensure emissions were lowered across the engine and vehicle operating range, Euro VI emissions legislation also includes Off Cycle Emissions (OCE) testing of operating conditions not featured in the WHC as part of the type approval process. The engine operating range is divided into grids and the OCE testing mandates measurement of randomly selected grids with Not to Exceed (NTE) limits legislated.

The reductions in emissions from Euro I to Euro IV legislations were largely met with reductions in 'engine out' emissions as a result of improved combustion. For FIS this resulted in a focus in improving the mixing of fuel and air in the combustion chamber, driving technological advancements in increased injection pressure and nozzle spray hole shaping. This was accompanied by a development of air-handling systems, including turbochargers and Exhaust Gas Recirculation (EGR).

To meet the Euro V emissions legislation, many OEMs had to further increase the complexity of air handling systems on the engine, increasing rates of EGR, and using turbochargers with Variable Geometry Turbines (VGTs) and increased intercooling. Euro VI also marked the introduction of EATS by some OEMs to control the legislated 'tailpipe' emissions. Both the further technological advances in air handling systems, and the introduction of EATS systems, necessitated parallel developments in FIS technology, to enable further increased injection pressures, variable injection rate shapes, and multiple injection strategies. Differing applications and manufacturers employ various combinations of combustion strategies at Euro V depending on the balance of air handling systems and EATS, requiring different injection strategies to be supported by FIS manufacturers, requiring flexibility in product characteristics.

Performance Stability Over Life

In addition to lowering the legislated emissions limits, Euro VI also included an extension to the emissions durability requirement for commercial vehicles and introduced In-Service Conformity (ISC) testing. From the introduction of the Euro IV legislation, manufacturers have been required to demonstrate that engines comply with emissions legislation, with an allowable deterioration relative to the legislation as-new limits, over a 'useful life', itself expressed as kilometre value as a function of vehicle class. As an example, a typical articulated haulage application would have to demonstrate compliance over a useful life of 500,000km. For Euro VI, that useful life was extended to be 700,000km, with proportionally similar increases for the other classes of commercial vehicles.

Prior to the introduction, the emissions durability of engines was determined by the removal of engines from their application and subsequent installation into an engine dynamometer facility. For Euro VI, a Portable Emissions Measurement System (PEMS) has been developed, allowing for ISC testing to be completed without removing the engine from the vehicle, allowing for emissions measurements to be taken in real life conditions and, as it is less disruptive than the previous method, to be conducted more frequently by the regulatory body. In the US market, the EPA have invested in a chassis dynamometer for commercial vehicles, allowing ISC testing to be completed with minimal preparation work required on the vehicle.

The extension of Emissions Durability requirements and enhanced ISC testing associated with Euro VI legislation mean that the performance stability of FIS is a significant issue for emissions compliance in application. This requires an FIS to operate reliably at the extremes of its operation range, necessitating stability in pressure at the nozzle and nozzle actuator performance over the engine's useful life. Significant changes in injector fuelling over life can also result in reduction of engine power and have the potential to turn on the engine warning light through detection by the On-Board Diagnostics (OBD) system.

Improved Vehicle Fuel Economy

Fuel efficiency and its laboratory equivalent CO_2 emissions, while not legislated in heavy duty markets, is a significant customer driver. CO_2 emissions of a vehicle are dependent on its design (including chassis layout, powertrain selection, and aerodynamics), its use environment and the specific duty cycle associated with its application. For passenger cars, the EU legislation assesses the vehicle's CO_2 performance through legislated drive cycles performed on a chassis dynamometer. While passenger car vehicles are relatively low in number of design configurations and have generally similar duty cycles, commercial vehicles exhibit significantly more diversity in design and application. For example, Mercedes-Benz report that they have in excess of 4500 possible configurations of truck available.

To legislate CO_2 testing via vehicle testing on chassis dynamometers would therefore place significant burden upon both the OEMs and the legislatures, in both the number of vehicles requiring testing, and the variety of test cycles required to assess different applications in a representative manner. As such, the current EC direction for future assessment and legislation of CO_2 emissions is through vehicle simulation which has been developed to be a means of providing reproducible, robust and practicable solution for generating realistic values that represent a commercial vehicle's real-world CO_2 emissions in application. The Euro VI legislation requires documentation of results to be provided by the manufacturer, but no limits or phased reductions are legislated.

The efficiency of an engine, reflecting on the efficiency of combustion, can be presented in terms of the Brake Thermal Efficiency (BTE), expressing the percentage of the thermal energy of the fuel that is converted to useful mechanical power. An increase in BTE represents reduced fuel consumption for a given power as the chemical energy in the fuel is converted into mechanical power more efficiently, and vice-versa. Historically, some legislated reductions in NO_x emissions have resulted in a drop in BTE that offset any improvements associated with air-handling and FIS. The focus for OEMs and their suppliers is now to enable BTE improvements once again through further technological development of the powertrain system, including FIS.

Reduced Noise, Vibration & Harshness

Noise, vibrations and harshness (NVH) has been recognised as increasingly significant in the commercial vehicle market. European legislation dictates a limit on vehicle noise of 80dB(A) measured at the end of a defined 2.5 second acceleration. Furthermore, the end users of the vehicles desire driving environments that are isolated from powertrain noise and vibrations to ensure they are comfortable working environments. Furthermore, the engines are required to be quick to start in a range of environmental conditions, and the operators of vehicles will consider any slow starting or increase in noise levels to be an indication of poor engine and FIS health. Implementation of engine stop-start systems in some medium duty applications place further significance on engine start performance. As such, OEMs require FIS systems that can contribute to reductions in NVH. FIS can contribute to reductions in noise as a result of pumping of fuel and as a result of combustion, while improvements to environmental stability can improve engine start-up performance.

Robust to Fuel Quality Variation

Fuel, in addition to being the medium being worked, acts as lubrication in FIS systems, and as such its composition and hydrodynamic bearing properties are of significance to FIS wear over life. The composition of the fuel, including its lubrication properties, can be shown to vary between relatively controlled markets, such as Europe and North America, and more dramatically in emerging markets such as Brazil, Russia, India, & China (BRIC). Furthermore, fuel composition can vary within a market as a function of the different additive packages employed by retailers, and seasonally in counties with extreme climates. Furthermore, there has been a continuing trend of legislative increases to the permissible blend ratio of Biodiesel in the fuel specifications for different markets, alongside a continued interest in the usage of increased blends of, including undiluted, Biodiesel and other alternative fuels. In addition, the quality of pump fuel, with respect to contamination levels of hard particles, water, and other non-hydrocarbon compounds, can have a significant detrimental effect on the performance, and wear over life, of FIS.

As fuel, both with regards to its composition and its quality, can be shown to be critical to the performance of FIS systems, and therefore the powertrain, over life, there has been a continued driver for diesel FIS systems to demonstrate a robustness to its variation.

System Integration and Packaging

Despite their physical size, many commercial vehicles are significantly constrained on the space available to package the powertrain. To enable modularity, the chassis and cab sections are separate constructions, and as a result, the powertrain, along with a large fuel tank for vehicle range, is largely constrained to the volume of the chassis section. As previously disucss, increasingly stringent emissions legislations have resulted in OEMs supplementing vehicle powertrains with multiple EATS systems. Such systems require significant packaging volumes for both the exhaust system itself (comprising both catalyst and filter elements), and in the case of SCR, for storage of the reagent (AdBlue or similar) and compressed air required for use. Furthermore, high-EGR emissions strategies place additional thermal management requirements on the vehicle, for both cooling of the EGR itself and often intercooling of the turbocharged intake system, resulting in a requirement to package additional heat exchangers. There is therefore an advantage associated with a FIS system that, in parallel with advances in air handling, can decrease engine-out emissions, reducing the reliance on EATS systems.

There is an additional drive from OEMs to enable existing engine architectures to be perpetuated for prolonged durations. The relative low volumes associated with heavy duty engines in comparison to passenger car engines, result in the desire for longer product life cycles to recover design, development and tooling costs. As such, base engine architectures are typically carried through multiple emissions legislations, enabled by design updates to the engine systems, including the addition of EATS systems. Therefore, there is an advantage associated to FIS technologies that enable performance improvements for existing engine platforms. For rapidly emissionising markets such as the BRIC markets, this allows OEMs to meet future emissions legislations at the lowest development costs through the implementation of state-of-the-art fuel injection systems.

Reduced Total Cost of Ownership

For the end user, typically owners and operators of single vehicles or larger fleets of vehicles, a significant driver behind their purchasing decisions is the Total Cost of Ownership (TCO) of a vehicle. The TCO of a vehicle considers the costs of the depreciation or leasing of the vehicle; hiring and paying a driver; servicing the vehicle; fuelling the vehicle; and purchasing other fluids (including engine oil and SCR Reagent). Typical annual TCO of a haulage vehicle will be in the region of €166k, for an annual drive distance of 120,000km.

While there is currently no legislative drive towards lowering fuel consumption and CO_2 emissions in the European commercial vehicle market, the distances associated with applications such as goods haulage, fuel costs represent a significant proportion of the operational costs for the end-users. In the past, relatively low fuel costs have resulted in fuel costs being as low as 10% of the annual cost of operating a commercial vehicle, but that has recently increased to as much as 30%. As such, there is increased emphasis from the end customer for commercial vehicles to demonstrate fuel consumption savings.

As such, there are significant benefits associated with any FIS system that can contribute to reductions in vehicle fuel consumption, through either increased operational efficiency or through enabling more fuel-efficient combustion. Fuel injection pumps are designed to achieve optimal volumetric efficiency at the speed and flow requirements typical of an application. The fuel injectors themselves are also designed to be as efficient as possible, through the reduction of parasitic leakage.

The purchase price of commercial vehicles, and therefore the proportion of TCO associated with depreciation or leasing of a vehicle, has increased as subsequent emissions legislations have required additional EATS systems to be added the powertrains. The challenge for OEMs is to balance the emissions performance of a vehicle system with its fuel efficiency and cost, both of development and for sale. Therefore, any additional improvements that can be made to 'engine out' emissions through advances in FIS technology can enable manufacturers to achieve the optimum system balance.

Appendix II Diesel fuel properties relevant to injector wear

The properties of diesel fuel that influence diesel FIS durability can be divided into three distinct groups. The first group includes kinematic viscosity and lubricity, the primary characteristics that affect wear and durability. The second group comprises parameters other than viscosity and lubricity that effect FIS system performance over life, such as sulphur and Fatty Acid Methyl Ester (FAME) concentrations. The third group concerns any contaminations to the fuel, including hard particles and ingress of fuel or lubrication oil. An additional contribution factor is associated with degradation of the fuel associated with aging and recirculation.

Group 1 – Kinematic Viscosity

The Kinematic Viscosity of a fuel is a measurement of its resistance to shear forces and flow. It is measured using a gravity flow meter at standardised temperatures and is typically expressed in values of mm2/sec at each temperature.

A low Kinematic Viscosity fuel used in high ambient temperatures can lead to difficulties with hot restarts or a loss of performance due to an inability to achieve pressure, while a fuel with an extremely low Kinematic Viscosity can result in a reduction in hydrodynamic and elasto-hydrodynamic lubrication. Furthermore, the viscosity of the fuel will influence the motion of valves, pins and nozzles through providing a damping force.

Group 1 – Lubricity

Lubricity is a measure of a fuel’s ability to reduce friction and wear in boundary lubricated, sliding interfaces. It is tested using a High Frequency Reciprocating Rig (HFRR) and measured through assessment of the resulting Wear Scar Diameter (WSD), measured in microns. A lower resultant WSD therefore indicates an increase in lubricity.

A low lubricity fuel can result in wear of boundary lubricated components within the injector, resulting in an increased propensity for failure of the FIS, and in the case of pumping elements, seizure of the plunger.

Group 2 – Initial Boiling Point

A fuel’s Initial Boiling Point (IBP) is the lowest temperature at which measurable evaporation occurs at atmospheric pressure. Diesel fuels can be manufactured to give a lower IBP through using lighter distillation fractions, but a low IBP is typically indicative of adulteration through the dilution of the diesel with kerosene or gasoline.

A lower IBP is conducive to cavitation occurring within the lower pressure regions of the injector, with the potential to cause cavitation erosion damage to surfaces and subsequent changes to injector performance. Furthermore, the high local temperatures resulting from the collapse of a cavitation bubble can also result in the formation of precipitates and deposits, resulting in a loss of injector performance and accelerated blocking of fuel filters.

Group 2 – 95% volume distillation temperature

The 95% Volume Distillation Temperature (95VDT) is the temperature at which 95% of the fuel will have evaporated. A high 95VDT is an indicator of a high concentration of non-volatile compounds within the fuel which do not combust and have a propensity to form deposits around the nozzle tip of injectors, reducing nozzle flows and engine power, and potentially resulting in a non-compliant emissions performance. As such, this failure mode is typically only assessed on combustion tests, such as engine and vehicle running, rather than non-fired validation tests and is sensitive to engine calibration.

Group 2 – Sulphur

All fuels contain sulphur compounds in varying concentrations, which contribute to exhaust particulates (SOx) and the poisoning of EATS systems. As a result, sulphur content is typically limited in many current diesel fuel specifications, with fuels having a concentration below 10mg/kg (Europe) and 15mg/kg (USA) being collectively referred to as Ultra Low Sulphur Diesel (ULSD). High sulphur content is therefore an indicator of a market with low quality fuel, unsuitable for advanced HPRCS systems, and can be observed at values up to 10,000mg/kg in some uncontrolled markets. Increased sulphur content can also be indicative of contamination by lubricant oils, particularly gear oils which often contain sulphur based Extreme Pressure (EP) additives.

Some sulphur compounds are actively corrosive, and in combination with water contamination can result in corrosive wear to fuel pumps and injectors, resulting in disruption in injection and potential failure.

Group 2 – Biodiesels (FAME)

Biodiesel, in the form of Fatty Acid Methyl Esters (FAME), has a lower stability than mineral diesel and degrades faster, resulting in the formation of organic acids and polymerisation compounds, in the form of gums and lacquers, in the fuel. Biodiesel content in pump fuel is regulated by concentration (limited to 7% by volume in typical developed fuel markets) and by Rancimat stability. The world markets have different standards for stability, with Europe legislating a blend stability of greater than 20hours, whereas the US market, which allows for a wider variety of biodiesel feed stocks, allows for blends to have a stability of less than 6 hours.

Current industry technology roadmap such as the Energy and Fuels Roadmap presented by the Automotive Council UK do not anticipate further increases in the blend ratio of FAME in diesel specifications, but do suggest increased usage of 'drop-in' Biodiesels such as Hydrogenated Vegetable Oil (HVO), that are transparent with fossil fuel diesel with respect to chemical properties, and are distributable through existing infrastructure, and would require no modifications to fuel systems or engines.

The use of FAME biodiesel in high blend concentrations and/or low stabilities can result in injector performance change or failure as a result of deposit formation associated with the acids and polymerisation compounds present in degraded fuel. FAME Biodiesel can also present issues with material compatibility, particularly in the case of seals, potentially resulting in external fuel leaks and a loss of injector performance.

Group 3 – Particle contamination

Adequate fuel cleanliness, particularly when concerning hard particle contamination, is an essential criterion to ensure durability of HPRCS systems. Many fuel standards specify a maximum allowable mass off contamination after a filtration process, such as the EN 12662 standard, while the ISO 11171 standard references both the size, and number of, particles present in fuel, in terms of particles larger than 4μm, 6μm and 14μm in size per ml of fuel. All Delphi Technologies FIS systems specify minimum fuel cleanliness at the inlet to the high-pressure system, after any fuel filters installed as part of the low-pressure system on the engine.

Hard particle debris contamination can result in abrasive wear to FIS components, particularly pumping elements and valve assemblies in HPCRS systems with tight tolerances, potentially resulting in a reduction in hydraulic efficiency of the FIS system, injector performance change over life, and rail pressure instability. In addition, particle contamination can result in accelerated fuel filter clogging, necessitating increased servicing requirements and reductions in system efficiency.

Particle contamination is a good overall indicator of fuel quality, where high levels of hard particle contamination correlate with the likelihood of other fuel quality issues. However, unlike many other fuel properties, particle contamination can be engineered for, through the use of appropriate on-engine filtration, servicing intervals, and low-pressure system design.

Group 3 – Elemental (inorganic) contamination

Concentrations of several inorganic elements, even in trace amounts, can potentially result in the formation of precipitates and deposits the FIS system, with Sodium, zinc, and copper being demonstrated to have adverse effects at concentrations as low as one part per million. Combinations of different trace elements, particularly zinc, calcium, and phosphorous, can be indicative of contamination by lubricant oil, either in oil lubricated fuel pumps, or through contamination of bulk storage tanks. Most fuel standards do not specify elemental concentrations.

The precipitates and deposits associated with elemental contamination are potentially harmful to FIS and EATS systems. Deposit formation within the injector control valve or nozzle can result in a change in injector performance, reducing power and affecting emissions compliance. Elemental contamination can also result in accelerated blocking of fuel filters, reducing system efficiency and increasing servicing requirements, while also potentially result in catalyst poisoning in the EATS systems.

Group 3 – Water content

Water can be present in the fuel system in one of three states: dissolved, emulsified, and as free water. Dissolved water can precipitate out of the fuel if it is cooled to saturation point. Undissolved water can result in corrosion of FIS components, a reduction in fuel stability (particularly in fuels with a biodiesel content), and accelerated growth of micro-organisms in bulk storage tanks with the potential to cause accelerated blocking of fuel filters. Water contamination typically occurs at the point of usage, but cases of fuel dilution at the point of distribution have been reported in uncontrolled fuel markets.

Fuel Degradation & Chemical Degradation

Diesel fuel can degrade in both storage and usage, adversely influencing its chemical and hydraulic properties. Oxidation of the fuel can result both from storage and through exposure of high temperature fuels to certain metals, to the detriment of its performance as a lubricant, while introducing precipitates of other chemicals that can increases the propensity of the fuel to form deposits in the FIS and the combustion chamber. On an engine or vehicle, the majority of the degraded fuel and any other chemical by-products are combusted, with only a small amount of fuel recirculating through FIS leakage, or the rail pressure control valve. However, on hydraulic test rigs, where it is typical to recirculate the fuel in order to reduce costs, fuel monitoring and replenishment should be considered. Fuel Degradation can be inferred through a number of metrics, including Total Acid Number (TAN).

Appendix III DoE applications in the automotive industry

This appendix presents a summary of the results of a review of SAE publications from 2000 onwards, with the keywords: “Design of Experiments”, “Factorial Design”, and “Taguchi Method”.

The abstracts of 167 papers were reviewed, some of which featured the application of more than one type of experimental design. Each paper was classified by experimental design type, and by application. The summary of that classification is as shown in Appendix III -1 and Appendix III-2.

Application Classification	**Count**	**Percentage**
CAE Design Optimisation	91	54.5%
Product Optimisation	37	22.2%
Combustion Optimisation	25	15.0%
Process Optimisation	9	5.4%
Failure Mode Characterisation	5	3.0%
Total	167	100.0%

Appendix III- 1 Summary of application classification

Design Classification	**Count**	**Percentage**
Taguchi Robust Design (TM)	86	46.5%
Factorial Design (FD)	36	19.5%
Fractional Factorial Design (FFD)	24	13.0%
Central Composite Design (CCD)	15	8.1%
Plackett Burman Design (PBD)	6	3.2%
Box Behnken Designs (BBD)	2	1.1%
Space Filling	16	8.6%
Total	185	100.0%

Appendix III- 2 Summary of Classification by Experimental Design

Appendix III-3 then provides a list of the 167 papers reviewed.

Paper	**Design 1**	**Design 2**	**Application**
Agudelo, Belcher & Dharaiya 2013	FFD		Product Optimisation
Akbarzadeh & Zohoor 2006	FFD		CAE Design Optimisation
Akehurst 2009	Space filling		CAE Design Optimisation
Amparo et al 2016	TM		Product Optimisation
Annabattula 2015	Space filling		CAE Design Optimisation
Apte & Hammoud 2005	FFD		Product Optimisation
Arrowsmith, Bott & Bush 2006	FD	CCD	Combustion Optimisation
Aseer et al 2017	TM		Process Optimisation

Avutapalli, Vallurupalli & Keshtkar 2003	TM		CAE Design Optimisation
Azadi & Mirzadeh 2005	FFD		CAE Design Optimisation
Bai et al 2015	CCD		Product Optimisation
Balaraman et al 2006	TM		Product Optimisation
Bayraktar et al 2005	FFD		CAE Design Optimisation
Beigmoradi 2015	TM		CAE Design Optimisation
Beigmoradi 2015	TM		CAE Design Optimisation
Belgiorno et al 2017	FD		Combustion Optimisation
Berntsson & Denbratt 2007	FD	FD	Combustion Optimisation
Birckett et al 2014	Space filling		CAE Design Optimisation
Biswas & Mandal 2013	TM		Product Optimisation
Brijesh, Chowdury & Sreedhara 2013	TM		Combustion Optimisation
Butkewitsch et al 2001	FFD	CCD	CAE Design Optimisation
Capetillo et al 2017	TM		CAE Design Optimisation
Chiang et al 2003	FFD		CAE Design Optimisation
Choi et al 2007	TM		CAE Design Optimisation
Daei, Davouddzadeh & Filip 2015	TM		CAE Design Optimisation
Dairou et al 2003	PBD		Product Optimisation
Dante, Carmen Fernadez & Rivacoba 2000	TM	TM	Product Optimisation
de Araujo & de Cunha Mello 2003	FFD		Failure mode
Demuynck et al 2912	FD		Combustion Optimisation
Dimitriou et al 2013	TM		Combustion Optimisation
Dukkipati, Srinivas & Mouli 2006	TM		CAE Design Optimisation
Eskandari, Mirzadeh & Azadi 2006	FFD		CAE Design Optimisation
Fang et al 2017	TM		Product Optimisation
Fernandes et al 2008	FD		CAE Design Optimisation
Flesch & Jaskulski 2005	FD		Product Optimisation
Flesch et al 2003	FD		Process Optimisation
Flesch, Ruschel & Tortorella 2004	FFD		Product Optimisation
Flesch, Zambarda & Silva 2003	FD		Product Optimisation
Fonseca, da Silva & Rossetti 2015	FD		CAE Design Optimisation
Freitas & Nonato 2010	FD	BBD	Product Optimisation
Fritz et al 2007	FD		CAE Design Optimisation
Ganti, Dewangan & Subramanian 2016	TM		CAE Design Optimisation
George, Chen & Shih 2013	FD		CAE Design Optimisation
Gilbert, Mandadapu & Cindric 2017	FD	TM	Product Optimisation
Gopalakrishnan et al 2014	Space filling		Combustion Optimisation
Gothekar et al 2007	TM		Combustion Optimisation
Govender & Barton 2009	TM		CAE Design Optimisation
Haenel et al 2011	TM		Combustion Optimisation
Hajireza et al 2006	FFD		CAE Design Optimisation
Hirsch & Re 2009	Space filling		CAE Design Optimisation
Hunt, Badarinarayan & Okamoto 2006	TM		Process Optimisation
Imoehl et al 2012	FD	FFD	Product Optimisation

Ioannou, Gurney & Downing 2005	Space filling		CAE Design Optimisation
Jadhav & Tandale 2015	TM		Process Optimisation
Jadhav 2015	TM		Process Optimisation
Jadhav 2016	TM		Combustion Optimisation
Jain et al 2016	TM		Combustion Optimisation
Jaradet et al 2008	BBD		Product Optimisation
Jasper, Rudell & Sakurai 2006	TM	FFD	CAE Design Optimisation
Jeon & Park 2007	FD		CAE Design Optimisation
Jian & Wang 2016	TM		CAE Design Optimisation
Jiang & Smith 2014	Space filling		CAE Design Optimisation
Jiang & Wang 2015	TM		CAE Design Optimisation
John et al 2013	TM		CAE Design Optimisation
Johnson, Kumar & Harne 2002	TM		CAE Design Optimisation
Jonkers, Bardon & Gardiner 2002	FD	FD	Combustion Optimisation
Junior & Galera 2015	FFD	PBD	CAE Design Optimisation
Kadam & Jadhav 2017	TM		Combustion Optimisation
Kawabe et al 2005	CCD		CAE Design Optimisation
Kawaguchi et al 2009	PBD	TM	CAE Design Optimisation
Keshavarz et al 2012	PBD		CAE Design Optimisation
Khaknejad & Keshavarz 2012	FFD		CAE Design Optimisation
Kim et al 2012	TM		Product Optimisation
Kim et al 2017	TM		CAE Design Optimisation
Kim, Cho & Yoon 2007	TM		CAE Design Optimisation
Kinchen 2017	TM		Product Optimisation
Kitamura et al 2011	Space filling		Combustion Optimisation
Koga et al 2011	TM		Failure mode
Kowada, Sato & Hosoi 2008	TM		Failure mode
Kumar et al 2002	TM		Product Optimisation
Kumar et al 2005	TM		CAE Design Optimisation
Kumar et al 2016	TM		CAE Design Optimisation
Kwon et al 2013	TM		CAE Design Optimisation
Landman et al 2002	FD		Product Optimisation
Le et al 2017	TM		Product Optimisation
Ledesma & Shih 2001	FD		CAE Design Optimisation
Lee & Hong 2003	TM		CAE Design Optimisation
Lee & Reitz 2003	FFD	FFD	Combustion Optimisation
Lee et al 2006	TM		CAE Design Optimisation
Lee et al 2009	TM		Product Optimisation
Lee, Ahn & Hong 2014	TM		CAE Design Optimisation
Leme et al 2014	FD		Product Optimisation
Li & Kim 2015	Space filling		CAE Design Optimisation
Lynch, Fleischmann & Duff 2002	FFD	CCD	Product Optimisation
Mallamo, Badami & Millo 2004	CCD		Combustion Optimisation
Manni & Florio 2013	FD		Product Optimisation

Millo et al 2007	FD		CAE Design Optimisation
Millo et al 2008	FFD		CAE Design Optimisation
Mittal, Gulve & Weaver 2006	CCD		CAE Design Optimisation
Mizuno et al 2012	TM		CAE Design Optimisation
Montgomery & Reitz 2000	FFD	FFD	Combustion Optimisation
Morgado 2004	FFD		CAE Design Optimisation
Mulani, Jadhav & Gopalakrishna 2008	TM		Product Optimisation
Murakami, Chinmoy & Asano 2009	TM		CAE Design Optimisation
Nagendiran, Sivanantham & Pandarinath 2011	TM		Product Optimisation
Nasser & Jawad 2008	CCD		CAE Design Optimisation
Nigade & Mutalikdesai 2016	TM		Process Optimisation
Nigade 2016	TM		Combustion Optimisation
Nishi, Rodrigues & Cardoso 2015	CCD		Combustion Optimisation
Novakova & Brun 2003	Space filling		CAE Design Optimisation
Okamura & Yumoto 2006	TM		CAE Design Optimisation
Pali, Kumar & Singh 2016	TM		Process Optimisation
Paliwal et al 2015	FD		CAE Design Optimisation
Paluskar & Vaidya 2011	TM		CAE Design Optimisation
Park & Yoon 2007	TM		Product Optimisation
Rahman, Ninawe & Solomon 2010	FD		CAE Design Optimisation
Rajamani, Schoenfeld & Dhongde 2010	FD		CAE Design Optimisation
Rao & Kumar 2013	Space filling		CAE Design Optimisation
Rathod & Rithe 2014	TM		Process Optimisation
Robere 2016	TM		Failure mode
Robinette & Singh 2016	TM		CAE Design Optimisation
Rose & Jebasinski 2003	TM		CAE Design Optimisation
Rossetti, Del Passo & Geremias 2013	FD		Failure mode
Roy & Landman 2006	FFD		Product Optimisation
Saha et al 2003	FFD		Product Optimisation
Saha, Haylett & Roy 2013	TM		Product Optimisation
Sarkar et al 2015	TM		CAE Design Optimisation
Sayyad, Salunke & Jadhav 2017	TM		Combustion Optimisation
Scheibner & Wendemuth 2009	CCD		Product Optimisation
Scherer 2017	TM		Product Optimisation
Schogl et al 2011	FD		CAE Design Optimisation
Schrangl et al 2016	Space filling		Combustion Optimisation
Scotto d'Apollonia, Granier & Debaty 2004	TM	CCD	CAE Design Optimisation
Seo 2011	TM		CAE Design Optimisation
Sert & Boyraz 2015	TM		CAE Design Optimisation
Sethuramalingam, Parmar & Tiwari 2016	TM		CAE Design Optimisation
Shankaranarayana 2017	FD		Product Optimisation
Sheridan & Wilson 2009	FD		CAE Design Optimisation
Shigarkanthi et al 2011	CCD		CAE Design Optimisation
Simoes et al 2001	PBD	RSM	Process Optimisation

Singh et al 2011	FD		Product Optimisation
Song & Tan 2011	TM		CAE Design Optimisation
Song, Lee & Ahn 2011	TM		CAE Design Optimisation
Soon, Gummadi & Cao 2005	TM		CAE Design Optimisation
Srinivasa, S, & Shome 2014	CCD		CAE Design Optimisation
Sriram, Speer & Matlock 1999	FD		CAE Design Optimisation
Stotera & Bombard 2015	TM		Product Optimisation
Suh, Lee & Yoon 2000	CCD		CAE Design Optimisation
Tankala, Marur & Wilson 2006	PBD		CAE Design Optimisation
Terra & Barbosa 2016	TM		CAE Design Optimisation
Waghmare et al 2017	Space filling		CAE Design Optimisation
Wang & Jiang 2015	TM		CAE Design Optimisation
Wang, Xu & Tan 2017	FD		Product Optimisation
Wang, Zhou & Xu 2017	TM		CAE Design Optimisation
Win et al 2002	TM		Combustion Optimisation
Xiaoyu, Guohua & Yankun 2004	FD		CAE Design Optimisation
Yalamanchili, Sharma & Thomson	Space filling		CAE Design Optimisation
Yamamoto et al 2002	TM		Combustion Optimisation
Yamaoka et al 2005	TM		Combustion Optimisation
Yang et al 2017	TM		CAE Design Optimisation
Yang, Men & Rowley 2008	FD	FD	CAE Design Optimisation
Yuan, Kane & Rahman 2016	Space filling		CAE Design Optimisation
Zhang et al 2008	Space filling	TM	CAE Design Optimisation
Zhang et al 2013	TM		CAE Design Optimisation
Zhou, Xu & Wang 2017	TM		CAE Design Optimisation
Zhu et al 2004	TM	CCD	Combustion Optimisation
Zutshi et al 2007	TM		CAE Design Optimisation

Appendix III- 3 List of papers reviewed

Appendix IV Questions asked to the experts in the Delphi Study

This appendix provides the questions asked to the expert panel in each round of the Delphi Study detailed in Chapter 6.

Round 1

1. Please can you define NCV seat wear?

   What variables (design/manufacture/usage) influence NCV seat wear?
2. Is it plausible and desirable for our products to be robust against NCV seat wear?
3. With respect to life wear out failure modes, is there a failure mode other than NCV seat wear that should be prioritised?

Round 2

The expert panel were provided with an editable table as shown in Appendix VI–1.

1. Review the variables listed, and add as you feel appropriate
2. Identify interactions between variables by placing variable numbers in the interaction column

   In the case below, I have (randomly) suggested that variables 2 & 29 are interacting as an example
3. Identify relative significance rankings for the variables based on your expert knowledge as being one of the following

   H = High

   L = Low

   N/A = Does not influence seat wear

   Variables left blank will be assumed to be of middling significance
4. Suggest, where appropriate, ‘levers’ [accelerating variables for testing purposes] that could be used to test the significance of each variable along with any future solutions
5. Please use the comments box for any justifications etc.

		Variable	Interactions	Relative ranking	Test 'levers'	Comments
Design	1	Material specifications (inc. coatings)				
	2	Seat geometries (diff. angle and contact area)	29			
	3	Geometric tolerances				
	4	Surface finishes				
	5	Concentricity tolerance				
	6	Tilt				
	7	Flow area/rates across seats				
	8	RDO positional tolerance				
	9	Seat profile waviness				
Manufacturing	11	Surface finishes				
	12	Deburr				
	13	Wash & cleanliness				
	14	Tool wear				
	15	Assembly of pin & armature				
	16	Clamp load distortion				
Usage	18	Number of injections				
	19	Injection durations				
	20	Operating pressure				
	21	Operating temperatures				
	22	Fuel lubricity				
	23	Fuel viscosity				
	24	Biofuel content				
	25	Fuel Initial Boiling Point				
	26	Hard particle debris contamination				
	27	Fuel deposits				
	28	Water-in fuel dilution				
	29	Oil-in-fuel dilution	2			
	30	Biofuel content degradation				
	31	Air ingress				
	32	Low pressure system backpressure				
	33	Pressure variation (dilation)				

Appendix IV - 1 Editable table for round 2 responses

Round 3

The expert panel were provided with an editable table as shown in Appendix VI–2.

1. Please review the variables listed as reaching consensus in round 2 (either as significant to NCV seat wear or not) and provide justified arguments should you disagree

   If your expert knowledge can help refine some of those variables, then please comment (as facilitator, I have not been applying any of my own 'expert' judgement in interpreting your responses, as such, some descriptors may seem coarse)
2. Please review the new variables identified in round two and indicate your judgement of their relative significance as before in the table below
3. For the variables where no consensus was met in round 2, please review the additional comments and justifications within the attached summary and indicate your judgement and justifications in the table below

Variable	Round 2 judgement	Relative ranking	Comments & justifications
Surface finishes			
Fuel viscosity			
Hard particle debris contamination			
Surface finishes			
Assembly of pin & armature			
Flow area/rates across seats			
Oil-in-fuel dilution			
Injection durations			
Operating temperatures			
Operating pressure			
Concentricity tolerance			
Pressure variation (dilation)			
Air ingress			
Low pressure system backpressure			
Wash & cleanliness			
Biofuel content degradation			
Fuel Initial Boiling Point			
NCV pin dynamic motion			
NCV lift			
Pin mass			
Pin stiffness			

Appendix IV - 2 Editable table for Round 3 responses

Round 4 (Optional)

1. NCV seat wear is not a new problem; what factors, be they internal or external, do you believe have restricted our capacity to fully characterise it in order to develop robust solutions?

Appendix V Publications

The following paper was accepted for the International Design Conference – Design 2018, in Dubrovnik, Croatia, May 21-24, 2018. However, the paper was not approved for release in time for final submissions, and as such, was not presented, nor featured in the conference proceedings. The paper predominantly forms a summary of the Delphi Study presented in Chapter 6, with appropriate elements from Chapters 3 and 9.

THE DELPHI METHOD AS A RESOURCE-SENSITIVE METHOD FOR EXPERT ELICITATION IN NEW PRODUCT DEVELOPMENT

PJ Bonnington[a,b,c], M Needham[a], CJ Brace[b], CA McMahon[c]

a. Delphi Powertrain Engineering, Stonehouse, UK
b. Department of Mechanical Engineering, University of Bath, Bath, UK
c. Faculty of Engineering, University of Bristol, Bristol, UK

Abstract

Failure mode characterisation is a key element of the New Product Development process, informed by expert judgment when insufficient codified knowledge is available. While existing elicitation methods are prevalent in the design process, few structured methods are prevalent in the design of product development tests. This paper presents the industrial application of the Delphi Method as part of a failure mode characterisation process, concluding that the method represents a resource sensitive approach for eliciting expert judgement in order to inform the design of structured experiments.

Uncertainty, Knowledge management, Decision making, Expert Judgement, Failure Mode Characterisation

Introduction

When dealing with uncertainty in product design & development programmes, ‘Expert Judgement’ describes the application of knowledge to inform a decision when objective data is unavailable or insufficient. Existing methods of eliciting expert judgement, such as Quality Function Deployment, are in common use in the product design process, but few such methods are prevalent in the design of empirical tests for the development and validation of products. With the appropriate experts distributed around an organisation, and failure mode knowledge required to design effective empirical programmes, there is a requirement to elicit expert judgement in a manner that is suitably flexible.

This paper presents an industrial study of expert judgement as an element of a failure mode characterisation process in the automotive industry, the results of which ultimately served as inputs for structured empirical investigation using Experimental Design methodology. The study took the form of an expert elicitation programme, utilising the Delphi Method, with a panel of technical experts from within the lead author's organisation. Furthermore, this study captures the challenges of conducting expert elicitation in a typical industrial environment.

While focused on a specific industry, this study is relevant to the wider context of organisations engaged in New Product Development (NPD) programmes, with distributed experts and a lean approach to product development, and presents the Delphi Method as a suitable, resource-sensitive approach to eliciting expert judgement to inform the product design and development process.

Background

The Commercial Vehicle Fuel Injection Systems (FIS) product line of Delphi Technologies is a UK-based product engineering group that designs, develops and validates FIS products for heavy duty applications. Delphi Technologies launched three ultra-high pressure common rail systems in 2012, utilising electronically actuated hydro-mechanical control valves with tolerances of single digit micrometres.

The three systems were developed by a multidisciplinary product engineering team across three different sites, using a product platform approach, resulting in commonality in design features and failure modes. The technology is advancing in response to the market drivers: a drive for engine efficiency is reflected in reducing parasitic leakage in the FIS; increases to robustness to wear and fuel quality are responsive to the market's need for high reliability; and increasing injection pressures ensure that Delphi Technologies' customers can meet increasingly stringent emissions legislation at the lowest system cost.

The resultant product characteristics are such as to increase the forces acting on the control valves, increasing the risk of mechanical wear. Furthermore, the highly controlled design tolerances increase the potential for the valves' performance over time to be sensitive to that wear.

While the wear of the control valve represents a relatively simple mechanical problem, it is embedded in a complex sociotechnical system. Through the parallel product development programmes, partial investigations have resulted in an incomplete characterisation of the problem. With the appropriate technical experts now engaged in subsequent customer-focused projects, a resource-sensitive approach was required to eliciting their expert judgement to inform a more rigorous empirical characterisation of the problem.

As in wider industry, expert judgement is used widely, and largely informally, within the Delphi Technologies' NPD process. Technical experts have been shown to have an important role in assessing the uncertainties encountered as part of a typical NPD programme (Bedford et al., 2006), and their opinions are considered essential when resources are limited, and when time-bound solutions to emerging problems are required (Burgman et al., 2011).

Appropriate use of expert judgement can provide decision makers with knowledge supplementary to objective data, and has been described in numerous contexts in literature. Goossens et al., (2013) suggest that expert judgement helps an organisation move towards a rational consensus from evidence in the event of uncertainty. It has been described as essential in the event of incomplete data, in unique contexts, or when extrapolation outside of existing knowledge is required (Burgman et al., 2011). In the context of NPD programmes, expert judgement is suggested as a means for assessing uncertainties that emerge during the design process (Bedford et al., 2006), and for structuring and parameterising a useful model in knowledge intensive processes (Ford & Sterman, 1998).

Expert judgement is elicited more formally through the product design process, with the most widespread example being the use of methods such as Quality Function Deployment in which the judgements of experts are called upon to augment codified knowledge, and to provide a substitute in the case of uncertainty. However, when it comes to the design of an appropriate empirical testing programme, we have not found evidence of similar widespread use. Empirical testing programmes associated with product validation, particularly reliability demonstration programmes, can be highly resource intensive, and if inappropriate design leads to the late discovery of new knowledge, it can be expensive to act upon. In order to best design the tests associated with product validation, knowledge of the failure mode(s) of the products is required (Nelson, 1990), and if uncertainty exists around that failure mode(s), expert judgement will be relied upon.

Expert Elicitation

Expert Elicitation (EE) describes the process(es) through which expert judgement is elicited, such that it can be incorporated into a decision-making process. EE can occur either formally or informally, but literature concurs that it represents a worthwhile tool in the decision-making process, and should be used frequently, regardless of formality (Kadane & Wolfson, 1998). In more informal applications, there are two roles: that of the decision maker who has the problem, and that of the experts that have the appropriate subject matter expertise (Bedford et al, 2006). In more formal applications, EE typically incorporates a third role, with the addition of an analyst responsible for identifying the appropriate experts, and for facilitating the elicitation process (ibid). Regardless of the method employed and experts selected, elicitation of expert judgement should be undertaken with the principles of: scrutability, fairness, neutrality, & performance control in mind (Goossens et al., 2008).

EE is subject to the influence of different forms of bias, both in the expert panel, and in the analyst alike. Three forms of possible bias in panels have been identified: motivational bias in the form of vested interests; anchoring bias where inappropriate prior data is transformed to suit the problem case; and availability bias where emphasis is placed on the most memorable, if not pertinent, data (Bedford et al, 2006). The analyst can introduce bias into the process through the sample data provided, particularly in quantitative studies (Kadane & Wolfson, 1998).

Studies have attempted to assess the reliability of judgements provided by individual experts through the use of paired comparisons (Goosens et al., 2008), with the potential to yield additional knowledge.

The Delphi Method

The Delphi Method (no connection to Delphi Technologies) is an established EE method which can trace its roots to Cold War-era US military contract work (Skulmoski & Hartman, 2007). The literature has shown that the method offers a flexible technique for exploring with a group of experts concepts both within, and outside of, the existing body of knowledge, while the method also allows for a flexible design with respect to both innovation and practicality, allowing the user to balance application with validity (ibid). The four key characteristics of the Delphi Method are: the anonymity offered to the panel; the iteration of opinion it encourages; the use of controlled feedback; and the aggregation of a group's response it enables (ibid).

Okoli & Pawlowski (2004) describe the Delphi Method as a means for best structuring communication such as to allow a group to deal with a complex problem, e.g. identifying variables, generalising theories, and understanding causal relations. The method has been identified as suitable for forecasting where knowledge is imperfect and group consensus is considered an acceptable alternative (Donohoe & Needham, 2009). Anonymity reduces the tendency for junior experts to align their views with senior colleagues, and can lower the probability of experts pursuing a tangential line of inquiry (Geist, 2010). Successful implementations have resulted in theory generation, and have been shown to derive new insight into the subject rather than just being a means of data collection (Day & Bobeva, 2005). It offers a more rigorous methodology for eliciting expert judgement than a traditional survey, offering flexibility in design, and leveraging an aggregated panel response to answer complex questions appropriately (Skulmoski & Hartman, 2007).

It is often mistakenly stated that the Delphi Method has the aim of reaching consensus, but instead the aim should be a stability in response, where a bipolar state of opinion amongst the expert panel is a valid and important result (Gordon & Pease, 2006). The value in the method is thus in the ideas it generates, through consensus and otherwise (ibid).

In the 50 years that the Delphi Method has been used in the public domain, numerous criticisms have been levied. Landeta (2006) presents a summary of the weaknesses of the Delphi, suggesting that one of the key characteristics of the method, its avoidance of face-to-face interaction, is also one of its most significant weaknesses, allowing experts to avoid direct challenges to their judgement. Additional methodological weaknesses include the time and effort required to complete a multi-round study, the ease with which the analysts' bias can influence the study, and the difficulty in confirming the accuracy and validity of the results (ibid). The literature has also shown that applications lacking rigour in design and panel selection can lead to unfavourable results (Skulmoski & Hartman, 2007)(Geist, 2010)(Landeta, 2006). Furthermore, the premise of the method assumes that the individual panel members are open to having their opinions challenged and changed by their peers (Donohoe & Needham, 2009). However, these limitations on validity are not unique to the Delphi Method, but instead are true of other qualitative group elicitation strategies (Geist, 2010).

The literature has identified a number of strategies considered as best practice for successful implementation of the method. It has been shown that the initial contact with the expert panel, in terms of the quality of the material provided and the engagement levels achieved, highly influences the effectiveness of a study. A formal invitation, detailing the research motives and the expert selection criteria has been shown to be effective in engaging with the expert panel (Donohoe 2009). Given the potential for availability constraints for typical experts, clarity of the time requirements for study participation is suggested to be provided upfront (Skulmoski & Hartman, 2007). In addition, it is recommended that the first round of the study should be possible to complete within 30 minutes (Okoli & Pawlowski, 2004).

Once initial panel engagement has been achieved, a number of strategies have been identified to improve panel response rate through panel retention, including ensuring a quick turnaround between rounds to maintain momentum, and limiting the duration of the panel to limit fatigue (Gill et al, 2013). Furthermore, it has been suggested that repeated communication with the panel members being described as 'experts' is a means for best ensuring continued engagement (Gordon & Pease, 2006).

Selection of the panel size is another consideration affecting both the potential results of the Delphi Method and the efficacy of its application. Large panels have the potential advantage of aggregating the judgement of a wider number of experts, but can be difficult to manage and can experience high attrition rates (Gill et al, 2013). Smaller panels, while easier to manage, can result in limited generalizability of the results, with bias in smaller number of experts influencing the response (ibid). The minimum panel size recommended by Donohoe & Needham (2009) is between 7 and 15while Okoli & Pawlowski (2004) recommend 10 to 18 members.

Designing the Delphi Study

This study forms an element of a wider research programme, seeking to develop a sustainable failure mode characterisation process in Product Validation, drawing from, amongst other elements, the knowledge embedded within the experts distributed through a business, to effectively characterise a problem, and design the appropriate empirical studies to explore robust design solutions.

Having been identified in Failure Mode and Effect Analysis and observed in use across multiple product platforms, the specific failure mode at the centre of this study (valve wear) had been previously investigated using several methods, including Fault Tree Analysis and 5 Whys. However, the prior studies had resulted in only partial characterisation. With the parallel and tangential investigations having been conducted by teams that were distributed geographically, in a project-based matrix organisational structure, there was the potential to have significant levels of uncodified knowledge in the business. This study was intended to combine that knowledge, such as to best structure further investigations.

This study had 3 main objectives. Firstly, group consensus on a definition of the problem was sought as a boundary for further investigation. Secondly, a detailed characterisation of the failure mode was sought through exploration of its influencing variables and their interactions in order to inform further empirical and analytical investigation. Finally, the panels worldviews, and their perceptions of the reasons for why Delphi Technologies had previously been unable to fully characterise the failure mode, were sought.

The technical experts within Delphi Technologies most suited for this study comprised of a number of individuals split across different geographical locations, supporting different projects, and occupying a range of roles within the organisation. Furthermore, many exhibited strong opinions and personalities, and previously, their opinions had not been very well conveyed, and/or received by their peers and superiors. The anonymity and avoidance of direct face-to-face interaction afforded by the Delphi Method was chosen to allow those experts to contribute to the study in the most effective manner.

The availability to attend physical meetings of geographically distributed technical experts with customer focused schedules can be greatly restricted (Geist, 2010). Indeed, in this study, the first available meeting opportunity with the entire expert group would have been limited to a 1 hour session, 5 weeks from the time of scheduling, while a more suitable half-day, co-located workshop would not have been possible to schedule within 6 months. In the context of this study, typical of many firms engaged in NPD, the distributed, remote engagement of the Delphi Method would allow for several iterative rounds of exploration to be completed before even the most restrictive of physical meetings could have been completed in this resource sensitive environment.

The application of the Delphi Method for this study has been designed with a view to the criticisms and lessons presented previously. A qualitative study was chosen to facilitate exploration and theory building around the characterisation of the failure mode. The study was conducted remotely via email, blindly distributed to the panel, with files attached by both the author/facilitator and the panel members alike. As several panel members shared the same open plan office space as each other and the author/facilitator, the panel were explicitly asked to not discuss the study directly with their colleagues, or the author/facilitator, in order to retain the desired anonymity.

For this study, the design of the Delphi Method was conceived with an notional framework, it was recognised that some evolution in design may be desirable. The application was designed to have 3 rounds, based on the 3 main objectives for the study, with the opportunity for the panel to review and reflect upon the group responses of the previous round. The panel members would then be able to iterate their responses in light of those group responses if their own judgement was influenced. The design intention was that the panellists would be able to complete each round in 30 minutes, but still allow sufficient scope for rich feedback as appropriate.

The first round of the study was designed as an opening round, with 3 short questions intended to promote engagement with the panel. The panellists were to be asked to provide their definition of the failure mode, identify any variables and their interactions that they judged to be of influence, and to establish their worldviews on the plausibility and desirability for the product to be robust to the failure mode, or whether they through there were alternative failure modes that should instead be prioritised.

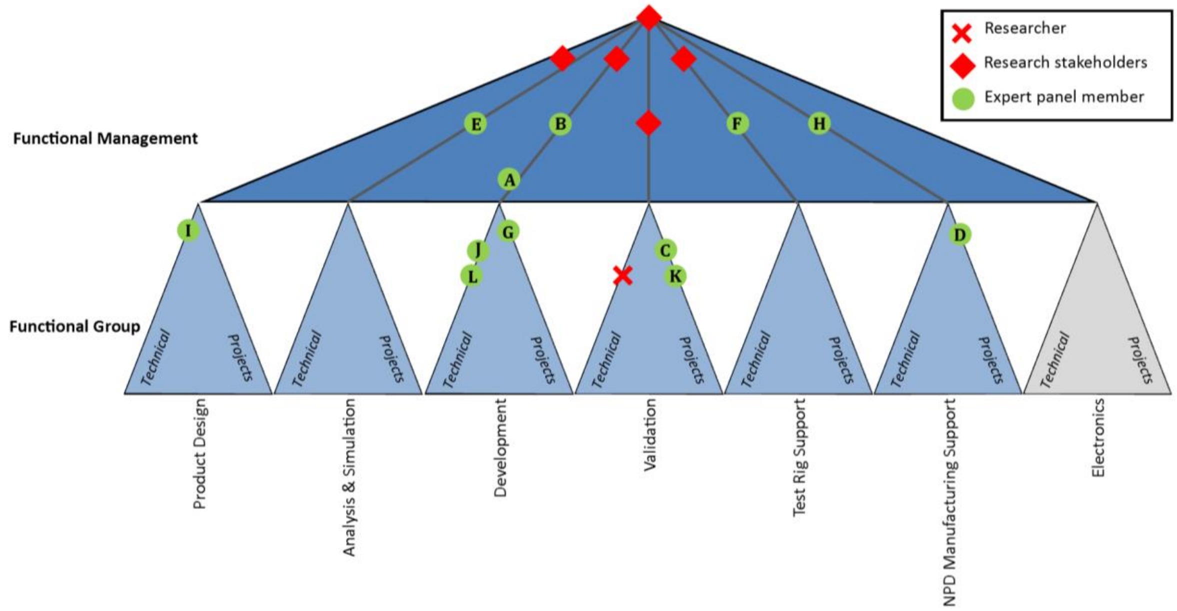
The design intention of the second round was to iterate the response to the first round in light of the group's collective response, before then exploring the relative perceived significance of the variables by asking the panel to weight each on a defined scale. The design intention of the third round was to once more to iterate the results of the previous round, before then asking the panellists to suggest the possible causes that have previously prevented Delphi Technologies from fully characterising this failure mode.

12 experts were invited to take part in this study, selected in conjunction with key research stakeholders, representing senior functional management, based on their expertise in the subject matter and their availability and motivation to take part in such a study alike, forming a homogeneous panel that met the minimum size recommendations identified in the literature. Subject matter expertise was defined as having direct experience in leading, or participating in, investigations into valve wear across any product platform past or present, either through product design, testing, or analysis. A number of the panellists were no longer directly involved in such investigations, or were no longer part of active product development functions, but were still identified has having relevant previous experience. No panel member had any prior experience with the Delphi Method, but all had participated in other forms of Expert Judgement Elicitation such as FMEA or 8D problem solving (Osma & Sayginer, 2010). The expert panel is summarised in Table 1, with their functional group at the time of the study, and relative positions and experience in the business identified. Figure 1 then visualises the relative position of each across the functional groups and hierarchy of Delphi Technologies, with both the researcher and the research stakeholders identified.

Functional groups and experience of Expert Panel

Expert	Functional Group	Seniority and experience in the business
A	Development	Team Leader, 10+ years
B	Development	Manager, 10+ years
C	Validation	Principal Engineer, 10+ years
D	NPD Manuf. Support	Principal Engineer, 20+ years
E	Analysis	Manager, 10+ years
F	Test Rig Support	Manager, 10+ years
G	Development	Principal Engineer, 10+ years
H	NPD Manuf. Support	Team Leader, 20+ years
I	Design	Principal Engineer, 15+ years
J	Development	Senior Engineer, 15+ years
K	Validation	Senior Engineer, 5+ years
L	Development	Senior Engineer, 10+ years

To facilitate both initial engagement and panel retention, several strategies previously identified in the literature were implemented for this study. As part of the invitation to participate, all panel members were provided with an information pack, detailing: the relevant background and motivation for the study; an overview of the Delphi Method; a timeline for completion of the study, including the dates associated with each round; and a terms of participation, outlining the expectations for the panel and the facilitator alike. The invitation took the form of an email providing a concise overview along the questions for round 1, with the information pack attached to prevent the panel being overloaded with material, allowing them to best consider the invitation to participate. The subject title of the email was chosen carefully to stand out from business as usual, and the logo of the supporting academic institution featured on all material supplied to the panel to enhance the creditability of the study.



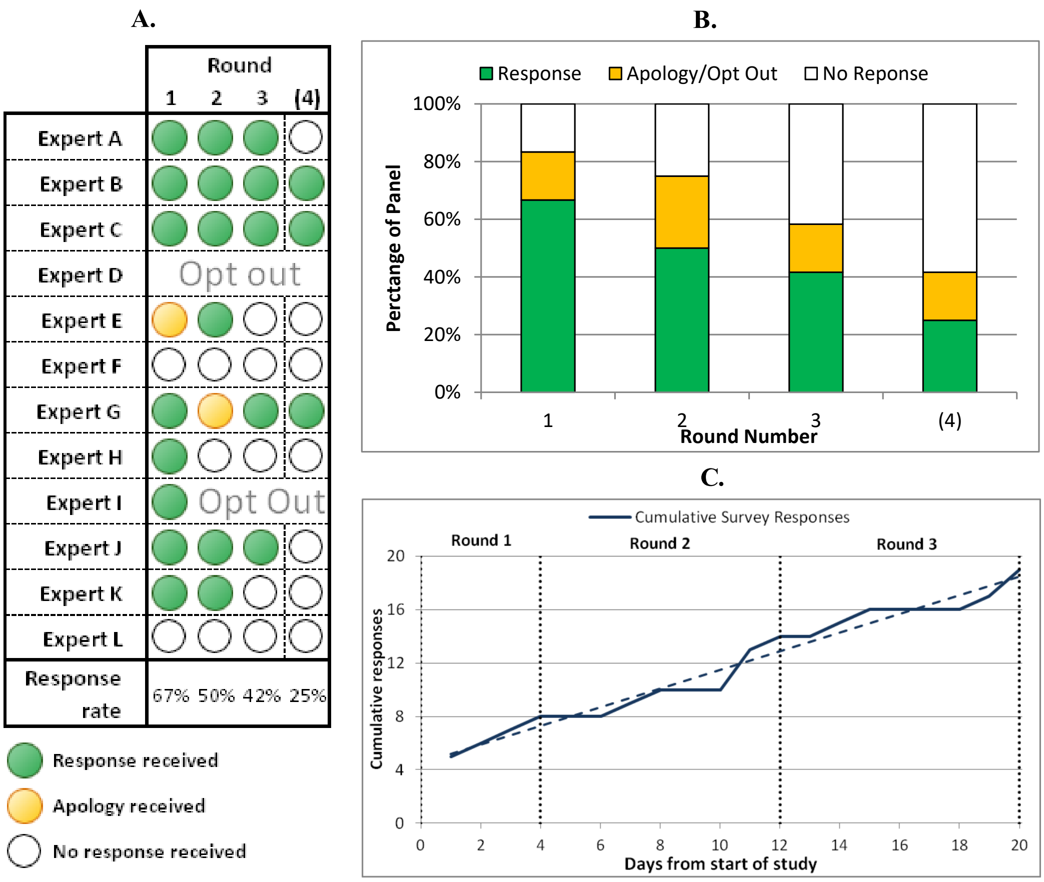
Relative functional position and seniority of each expert panel member

In order to limit the time and effort required to complete the study, effort was made to limit the scope of each round, and to provide template for responses, such that each round could be completed in 30 minutes or less, as outlined in the agreed terms of participation. To best ensure that the study adhered to the agreed timeline of the programme, panel members were reminded before the close of each round, with a 1-day window for late submissions to their availability. The results of the previous round, along with any new questions, were then provided within 2 working days, adhering to the terms of participation whilst reflecting panel the retention strategy.

Executing the Delphi Study - A reflection

The 3-round structure of the study as outlined in the terms of participation was adhered to, but the specific structure of each round required some evolution in order to derive best value from the study. While rounds 1 & 2 remained largely as designed, a decision was made to alter the focus for round 3 as a result of a combination of emergent factors. Firstly, the ranking of variables completed in round 2 showed significant disagreement between the expert panellists. Secondly, the panel had exhibited a consistently low willingness to reflect upon, and iterate the results of previous rounds, instead focusing on progressing with any new questions, to the detriment of the iterative nature of the Delphi Method. This is either indicative of a flaw in the methodology or the application, or perhaps a symptom of the environment the expert panel are used to working in.

In addition, several members of the panel did not provide justifications to their rankings provided in round 2, potentially making it difficult for the remainder of the group to reflect upon with respect to their own judgement. As such, round 3 was re-designed to explicitly task the expert panel to reflect upon and re-rank the disputed variables identified in round 2 in light of the justifications that were provided by the panel. This lack of justifications could indicate that either the panel members were unsure of their opinions, or that they were time-bound in their response. The question that was originally designed to be asked in round 3 was instead asked as an optional closing question. This provided the panel an option to participate in a fourth round without deviating from the agreed terms of participation. The panel's responses were tracked through the study to allow reflection on the application of the Delphi Method. Figure 2.A shows the responses received from each panel member for each round, including the optional round 4. No nominations of additional panellists were received as a result of the initial invitation, and one panel member chose to opt-out from the study citing a perceived lack of knowledge of the failure mode in application. A further expert then chose to opt-out after round 1, citing the results presented as being inappropriately 'too vague or too specific', and while he was reminded of the iterative nature of the Delphi Method, and the opportunity it would present to better characterise the failure mode with his peers, he elected to take no further part. As can be seen, a combination of reasons, including annual leave and specific customer deadlines, led to varying participation levels within the panel, providing evidence of the limitations of EE in resource sensitive environments.



Visual summary of study participation

Figure 2.B shows the response rate for each round, which can be seen to decline in an approximately linear fashion. This observation could be indicative of either panel fatigue, or the more general issues associated with the availability of experts within Delphi Technologies. While the panel reduced in size to below that recommended in literature , Donohoe & Needham (2009) did note that homogeneous panels can be smaller.

Figure 2.C then shows the cumulative responses for each working day for which the study was open. As can be seen, by the close of the first working day, 5 responses had been received demonstrating that the study generated high interest and engagement in the experts within Delphi Technologies.

The anonymity of the study was partially breached through discussions of the study in an open-plan office space, and through some experts suggesting they could identify fellow panellists through their responses, highlighting a potential limitation of an internal, homogenous panel.

Results of the Study

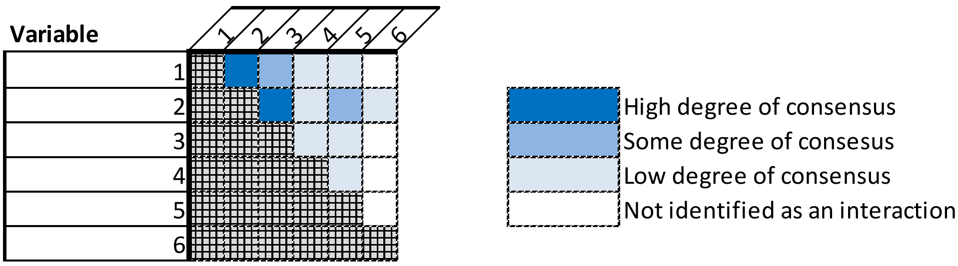
Group led definition

In round 1 of the study, the panellists were asked to define the failure mode. The individual panel members were asked to provide a free text response that would then be transformed into an aggregated response through qualitative data analysis. The author/facilitator assessed the responses, capturing any recurring themes and identifying the emerging consensus. The result included then sentence fragments provided directly by the panel when appropriate. The proposed definition was restrictive in that it referenced sliding wear of the valve seats only, with no reference to wear from impact or erosion.

In round 2, the experts were asked to review this definition, and either approve or suggest changes as appropriate. No changes were suggested by the panel and the definition was approved.

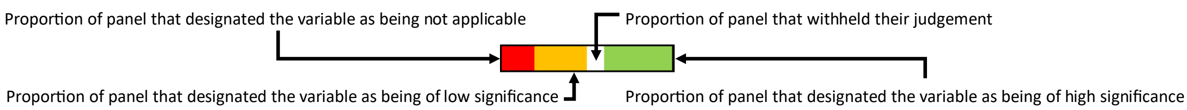
Variables & interactions

When asked in round 1 to identify the variables that influenced Control Valve seat wear, the panellists identified a total of 31 variables, grouped by them as either pertaining to the design, manufacture, or usage of the Control Valve. The identification of unique variables required an element of judgement from the author/facilitator to identify instances where different terminology was used by panellists, highlighting the requirement for an agreed taxonomy. The grouped variables were presented in the feedback to the panel, and no comments, clarifications, or changes were suggested by the panel.
In round 2, the experts were asked to explore the interactions between the variables that were identified in round 1. Markedly different levels of resolution were employed by the experts during this process, with some focusing only on the highest level interactions, listing 10 interactions, while others identified every action regardless of perceived significance, listing over 100 interactions.
The reliability of the experts was assessed by their consistency in correctly labelling pairs of interacting variables, in a similar manner to the 'paired comparisons' metric employed in literature. Only one expert (B) made any mistakes, failing to correctly label 15 paired interactions, but given that they identified significantly more interactions than any other expert, this only resulted in a 7% error rate.
The group results were aggregated using matrix addition for the variable pairs identified by each expert. The resultant matrix was such that any the higher the value for each variable pair, the greater the level of consensus of its significance to Control Valve seat wear. This was summarised for the feedback to the panel using a colour gradient, an example of which is shown in Figure 3.



Example of feedback for variable interactions

In round 2, the panel was also asked to assess the significance of the variables identified in round 1, reviewing review the variables, along with any justification or evidence supplied by their peers, and rank each one as being of either high or low significance to Control Valve seat wear, to be not applicable, or to withhold their judgement. While some experts assessed the significance of each variable, others elected to only rank those that they considered as either of high significance, or as not applicable, suggesting that they chose to withhold their judgement on the variables they were less certain of, or that they only wished to save time in responding by only ranking the most and least applicable variables. A total of 4 new variables were also identified by panellists in round 2, and were presented to the panel for their further review.
In addition, the experts were asked to provide justification for their judgements to share with the panel. However, many experts provided no justification or evidence to support their rankings at this time.
The results of this ranking process were presented back to the panel using an adapted 'Italian Flag' notation (Blockley & Godfrey, 2000), identifying the degree of consensus for each variable as shown in Figure 4. The author/facilitator then pooled the variables into 4 groups: those with a consensus as being of high significance; those with a consensus as being of little to no significance; new variables identified in round 2; and those were contradicting judgements were witnessed. Despite one expert electing to opt out after round 1 as the variables identified were either too specific, or too vague, no refinements to the definitions, nor suggestions for either splitting or merging variables, were received.



Notation used to feedback panel judgements

In round 3, the panel was asked to review the rankings and evidence for the 17 round 2 variables where no consensus was met, and the 4 new variables identified in round 2, and iterate their own judgements of their significance. The panel were supplied with a summary of the anonymised, verbatim rankings and justifications provided by their peers in round 2, along with the interactions identified for reference. The panel was asked to reflect on this summary, and provide a second ranking with revisions and justifications as appropriate. The proportion of experts that provided qualitative justifications alongside their rankings increased in round 3 although some experts still failed to provide any comments alongside their decisions.

When presented with the evidence and justifications supplied by their peers, the majority of the active panel members chose to revise some of their own judgements, with the exception of one panellist who elected to make no changes to his previous judgements, nor did he provide any justifications in doing so. Of those experts who did elected to change their own judgements: one acknowledged that he lacked evidence to support their previous judgement; one agreed that some variables were open to differing interpretations; and others provided no justifications at all.

The results of this iteration were summarised through a visualisation of the relative consensus before and after round three, alongside a comment from the author/facilitator on the degree of any change in consensus, an example of which is shown in Figure 5. The majority of the 17 previously discordant variables showed a degree of change in relative consensus after round 3, but only 4 moved to consensus as either being of high or little to no significance to the failure mode. No consensus was reached on the 4 new variables identified in round 2, suggesting further exploration may be appropriate as completed for the original variables.

Variable	Round 2 judgement	Round 3 judgement	Summary of Round 3
8			Several revisions towards consensus as significant

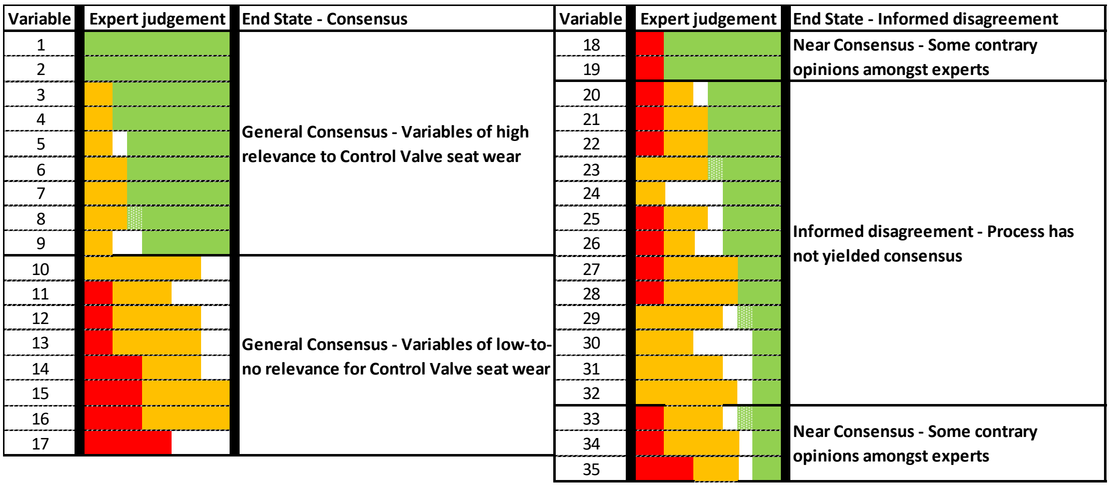
Summary of round 3 showing any change in group consensus

Three variables almost reached consensus as being of little to no significance to Control Valve seat wear. Given the justifications provided by experts on both sides of the argument, there is evidence of anchoring bias in some panellists.

After round 3, 13 variables, including the variables added in round 2, showed no consensus in the judgement of their significance amongst the expert panel. Given that the variables had been reviewed by the panel, with evidence and justifications provided by the experts, with no consensus being met, these variables can be described as having reached informed disagreement on their significance.

Throughout this process, only one expert referred their judgements on variables back to the group-led definition which referenced only sliding wear of the seating components, with no inclusion of either impact wear or flow erosion. However, several variables identified by the panel were either solely, or significantly related to impact or flow erosion rather than sliding wear. This suggested that the group-led definition perhaps required further iteration as the panel had identified variables deemed of significance that were not associated with the wear mode it described.

A final summary of the 35 variables was presented to the panellists, with the variables grouped by degree of consensus reached, as shown in Figure 6. In addition, for each variable where informed disagreement was reached, a summary of all evidence and justifications was collated as an output of the study.



Summary of Delphi Study

Worldviews on the feasibility and desirability of robustness against seat wear

In order to capture the world views of the expert panel members, in round 1 they were asked to comment on their perception of the feasibility and plausibility of Delphi Technologies' products being robust against Control Valve seat wear. The panel’s responses were consistent in their response that it is desirable for Delphi Technologies’ products to be robust against control valve seat wear, but the group response suggested that it is not plausible to be 100% robust against seat wear given the current design concept. Instead, the group suggested that engineering effort should focus on reducing it to the point of insignificance with respect to the life of the product through the identification of the root cause(s) and implementation of appropriate design solutions. The group also agreed that an alternative solution would be to reduce the products sensitivity to control valve seat wear, even if this was to the detriment of as-new performance.

Despite the panel members representing a wide cross section of functional groups, levels of seniority, and physical locations, the world views of the panel on the desirability and plausibility of improving robustness to Control Valve seat wear were consistent, with no disagreements identified.

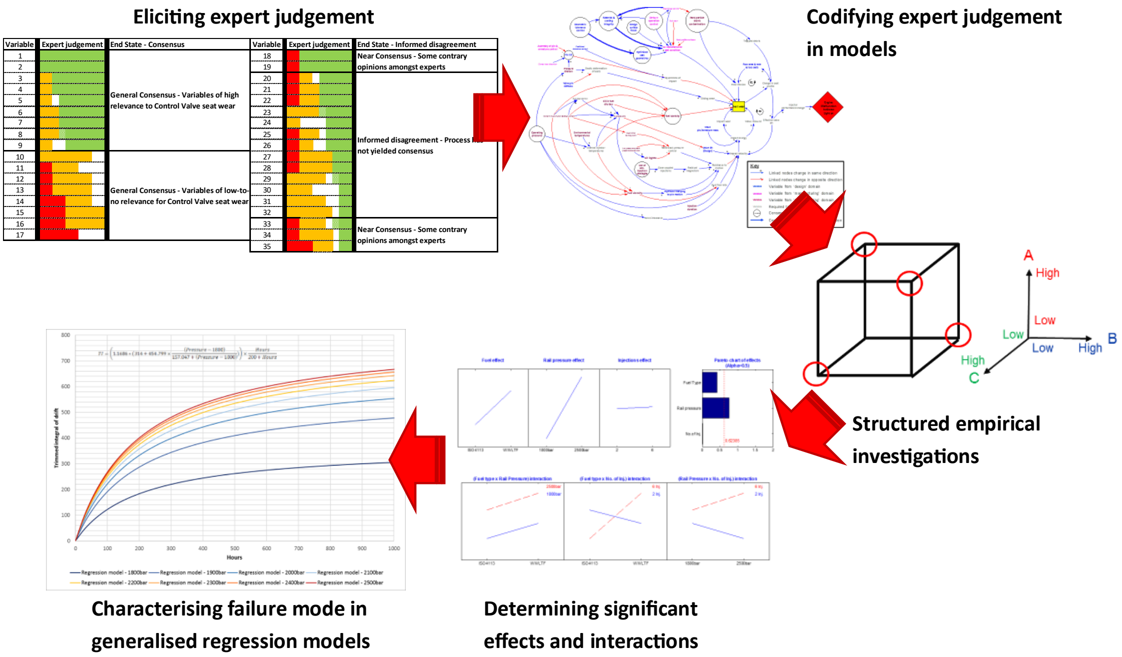
The panel were also asked to identify any through-life failure modes that they identified as being of higher priority to the business. Panel members identified a total of 4 wear-out failure modes judged to be significant, 2 of which being identified by multiple experts, but agreed that none were of higher priority than Control Valve seat wear at the time of the study.

Exploring the soft system causality

As the study was redesigned to encourage further iteration of judgement in the characterisation of the failure, the panellists were provided with the option to participate in a fourth round to explore the final objective of the study. Three panellists opted to take part, providing free text responses for what they judged as the reasons why Delphi Technologies had been previously unable to fully characterise this failure mode. While participation in this round was low, the quality and quantity of the qualitative feedback was markedly improved, and was then analysed by the author/facilitator for recurring themes and causal links, before then construction a System Dynamics model of the system. This represents continuing work and is outside the scope of this paper.

The Delphi Method as an element of a systematic approach to failure mode characterisation

The Delphi study presented in this paper is an element of a wider research programme, in which a failure mode of Diesel Fuel Injectors is characterised into generalised regression models that are in turn validated with existing data. Future publications will provide additional detail of the remaining elements of the failure mode characterisation process, but a visualisation of the approach is shown in Figure 7. After codification in a system model of the failure mode, the expert judgement elicited in this study served as the basis for an experimental investigation in which the variables identified as being of high significance to the failure mode, and those where informed disagreement was met, were considered as potential design factors. The characterisation process resulted in generalised regression models of the failure mode with respect to usage variables, providing a controlled and repeatable test for robust design solutions.



The Delphi Method as an element of failure mode characterisation

Conclusions

This study has demonstrated the effectiveness of the Delphi Method in a theory building mode, in a mature organisation engaged in NPD, providing an effective means for eliciting expert judgement into the product development process in a resource sensitive manner. It has been demonstrated that with the selection of a relevant subject, an appropriate panel, and an engagement strategy, the Delphi Method can be effective in involving distributed, customer focused experts, with high initial participation and response speed. However, the linear decrease in response rate in this study provides evidence on the difficulties associated with expert elicitation.

The expert panel demonstrated a low focus on reflection & on the provision of rich qualitative feedback, either a symptom of the sensitivity associated with the availability of this resource, or of a conditioning associated with dealing with emergent, customer driven problems. However, the flexibility of the Delphi Method allowed the design of this study to evolve to accommodate that.

One of the key motivations for the selection of the Delphi Method in this study was the anonymity it affords to the expert panel. This study demonstrated a potential limitation when applied to an internal, partly homogeneous panel, with partial breaches observed.

In this study, the application of the Delphi Method resulted in the characterisation of a failure mode agreed by the expert panel to be the most significant currently witnessed. A total of 35 variables, and their interactions, associated with the design, manufacture, and usage of the products, were identified as influencing the failure mode, with significances attributed to each. Through the iteration of judgement associated with the Delphi Method, the value of group expert judgement over that of single technical experts was demonstrated through the individual identification of 11 variables which the panel iterated to then judge as of little to no significance to the severity of the failure mode. The informed disagreement associated with 14 variables witnessed in this study is as important a result as for those where consensus was met, presenting areas suitable for future investigations.

As part of wider research, this study has demonstrated the effectiveness of the Delphi Method as part of an iterative, reflective, and resource sensitive approach to failure mode characterisation in NPD.

Acknowledgement

- The lead author would like to acknowledge his colleagues at Delphi Technologies for their continued support, and the EPRSC for funding this research.

References

Bedford, T.J., Quigley, J.L. & Walls, L. a. (2006), "Expert elicitation for reliable system design", Statistical Science, Vol. 21 no. 4, pp.428–450. doi: 10.1214/088342306000000510.

Blockley, DI & Godfrey, PS (2000), Doing it Differently, Thomas Telford, London. doi: 10.1680/didsfrc.27480

Burgman, M.A., McBride, M., Ashton, R., Speirs-Bridge, A., Flander, L., Wintle, B., Fidler, F., Rumpff, L. & Twardy, C.(2011), "Expert status and performance", PLoS ONE, Vol. 6 No. 7, pp.1–7. doi: 10.1371/journal.pone.0022998

Day, J. & Bobeva, M. (2005), "A Generic Toolkit for the Successful Management of Delphi Studies", Electronic Journal of Business Research Methods, Vol. 3 No. 2, pp.103–116.

Donohoe, H.M. & Needham, R.D. (2009), "Moving Best Practice Forward: Delphi Characteristics, Advantages, Potential Problems and Solutions", International Journal of Tourism Research, Vol 11, , pp.415–437. doi: 10.1002/jtr.709

Ford, D. & Sterman, J.D. (1998), "Expert Knowledge Elicitation to improve formal and mental models", System Dynamics Review, Vol. 14 No. 4, pp.309–340. doi: 10.1002/(sici)1099-1727(199824)14:4%3C309::aid-sdr154%3E3.0.co;2-5

Geist, M.R. (2010), "Using the Delphi method to engage stakeholders: A comparison of two studies", Evaluation and Program Planning, Vol. 33 No. 2, pp.147–154. doi: 10.1016/j.evalprogplan.2009.06.006

Gill, F.J., Leslie, G.D., Grech, C. & Latour, J.M.(2013), "Using a web-based survey tool to undertake a Delphi study: Application for nurse education research", YNEDT, Vol. 33 No. 11, pp.1322–1328. doi: 10.1016/j.nedt.2013.02.016

Goossens, L.H., Cooke, R.M., Hale, A.R. and Rodić-Wiersma, L. (2008), "Fifteen years of expert judgement at TUDelft", Safety Science, Vol. 46, pp.234–244. doi: 10.1016/j.ssci.2007.03.002

Gordon, T. & Pease, A. (2006), "RT Delphi: An efficient, “round-less” almost real time Delphi method", Technological Forecasting and Social Change, Vol. 73 No. 4, pp.321–333. doi: 10.1016/j.techfore.2005.09.005

Kadane, J.B. & Wolfson, L.J. (1998), "Experiences in elicitation", The Statistician, Vol. 47 No. 1, pp.3–19. doi: 10.1111/1467-9884.00113

Landeta, J. (2006), "Current validity of the Delphi method in social sciences", Technological Forecasting and Social Change, Vol. 73, pp.467–482. doi: 10.1016/j.techfore.2005.09.002

Nelson, W. (1990), Accelerated Testing: Statistical Models, Test Plans and Data Analyses, John Wiley & Sons. Inc., New York. doi: 10.2307/1269050

Okoli, C. & Pawlowski, S.D. (2004), "The Delphi method as a research tool: an example, design considerations and applications", Information & Management, Vol. 42 No. 1, pp.15–29. doi: 10.1016/j.im.2003.11.002

Osma, A., & Sayginer, A. S. (2010), "An Assessment on Problem Solving Approaches in Automotive Industry: Illustrative Templates, Similarities and Differences", Proceedings of the SAE World Congress. doi: 10.4271/2010-01-0698

Skulmoski, G.J. & Hartman, F.T. (2007). "The Delphi Method for Graduate Research". Journal of Information Technology Education, Vol. 6 No. 1, pp.1–21.